Multivariable and Vector Calculus

David A. SANTOS dsantos@ccp.edu

Mathesis iuvenes tentare rerum qualibet ardua semitasque non usitatas pandere docet.

July 17, 2008 Version

Preface

These notes started during the Spring of 2003. They are meant to be a gentle introduction to multivariable and vector calculus.

Throughout these notes I use Maple™ version 10 commands in order to illustrate some points of the theory.

I would appreciate any comments, suggestions, corrections, etc., which can be addressed to the email below.

David A. SANTOS
dsantos@ccp.edu

1 Vectors and Parametric Curves

1.1 Points and Vectors on the Plane

We start with a naïve introduction to some linear algebra necessary for the course. Those interested in more formal treatments can profit by reading [BlRo] or [Lan].

1 Definition (Scalar, Point, Bi-point, Vector) A *scalar* $\alpha \in \mathbb{R}$ is simply a real number. A *point* $\mathbf{r} \in \mathbb{R}^2$ is an ordered pair of real numbers, $\mathbf{r} = (x, y)$ with $x \in \mathbb{R}$ and $y \in \mathbb{R}$. Here the first coordinate x stipulates the location on the horizontal axis and the second coordinate y stipulates the location on the vertical axis. See figure 1.1. We will always denote the origin, that is, the point $(0, 0)$ by $\mathbf{O} = (0, 0)$. Given two points $\mathbf{r}$ and $\mathbf{r}'$ in $\mathbb{R}^2$ the directed line segment with departure point $\mathbf{r}$ and arrival point $\mathbf{r}'$ is called the *bi-point* $\mathbf{r}$, $\mathbf{r}'$[1] and is denoted by $[\mathbf{r}, \mathbf{r}']$. See figure 1.2 for an example. The bi-point $[\mathbf{r}, \mathbf{r}']$ can be thus interpreted as an arrow starting at $\mathbf{r}$ and finishing, with the arrow tip, at $\mathbf{r}'$. We say that $\mathbf{r}$ is the *tail* of the bi-point $[\mathbf{r}, \mathbf{r}']$ and that $\mathbf{r}'$ is its *head*. A *vector* $\overrightarrow{\mathbf{a}} \in \mathbb{R}^2$ is a codification of movement of a bi-point: given the bi-point $[\mathbf{r}, \mathbf{r}']$, we associate to it the vector $\overrightarrow{\mathbf{rr}'} = \begin{bmatrix} x' - x \\ y' - y \end{bmatrix}$ stipulating a movement of $x' - x$ units from (x, y) in the horizontal axis and of $y' - y$ units from the current position in the vertical axis. The *zero vector* $\overrightarrow{\mathbf{0}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ indicates no movement in either direction.

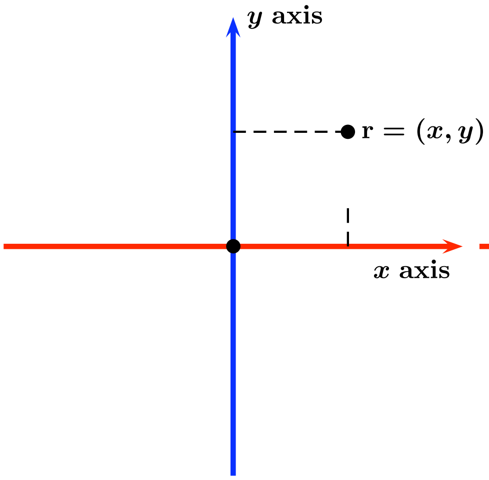

Figure 1.1: A point in $\mathbb{R}^2$.

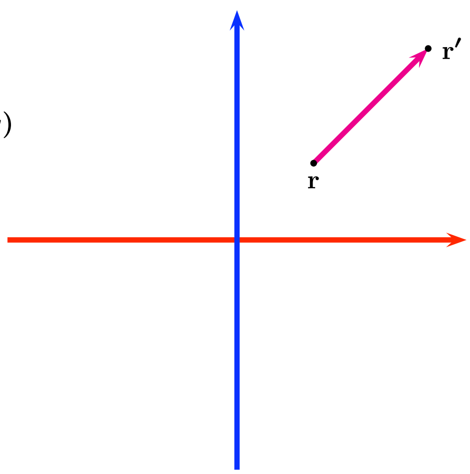

Figure 1.2: A bi-point in $\mathbb{R}^2$.

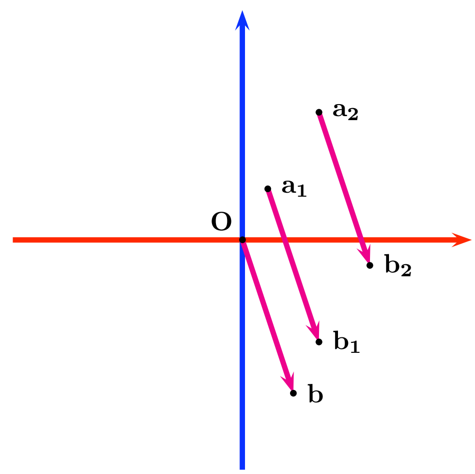

Figure 1.3: Example 2.

Notice that infinitely many different choices of departure and arrival points may give the same vector.

2 Example Consider the points

$$\mathbf{a}_1 = (1, 2), \quad \mathbf{b}_1 = (3, -4), \quad \mathbf{a}_2 = (3, 5), \quad \mathbf{b}_2 = (5, -1), \quad \mathbf{O} = (0, 0) \quad \mathbf{b} = (2, -6).$$

[1] Some authors use the terminology "*fixed vector*" instead of "bi-point."

Though the bi-points $[\mathbf{a}_1, \mathbf{b}_1]$, $[\mathbf{a}_2, \mathbf{b}_2]$ and $[\mathbf{O}, \mathbf{b}]$ are in different locations on the plane, they represent the same vector, as

$$\begin{bmatrix} 3-1 \\ -4-2 \end{bmatrix} = \begin{bmatrix} 5-3 \\ -1-5 \end{bmatrix} = \begin{bmatrix} 2-0 \\ -6-0 \end{bmatrix} = \begin{bmatrix} 2 \\ -6 \end{bmatrix}.$$

The instructions given by the vector are all the same: start at the point, go two units right and six units down. See figure 1.3.

In more technical language, a vector is an *equivalence class* of bi-points, that is, all bi-points that have the same length, have the same direction, and point in the same sense are equivalent, and the name of this equivalence is a *vector*. As an simple example of an equivalence class, consider the set of integers $\mathbb{Z}$. According to their remainder upon division by 3, each integer belongs to one of the three sets

$$3\mathbb{Z} = \{\ldots, -6, -3, 0, 3, 6, \ldots\}, \quad 3\mathbb{Z}+1 = \{\ldots, -5, -2, 1, 4, 7, \ldots\}, \quad 3\mathbb{Z}+2 = \{\ldots, -4, -1, 2, 5, 8, \ldots\}.$$

The equivalence class $3\mathbb{Z}$ comprises the integers divisible by 3, and for example, $-18 \in 3\mathbb{Z}$. Analogously, in example 2, the bi-point $[\mathbf{a}_1, \mathbf{b}_1]$ belongs to the equivalence class $\begin{bmatrix} 2 \\ -6 \end{bmatrix}$, that is, $[\mathbf{a}_1, \mathbf{b}_1] \in \begin{bmatrix} 2 \\ -6 \end{bmatrix}$.

3 Definition The vector $\overrightarrow{\mathbf{Oa}}$ that corresponds to the point $\mathbf{a} \in \mathbb{R}^2$ is called the *position vector* of the point $\mathbf{a}$.

4 Definition Let $\mathbf{a} \neq \mathbf{b}$ be points on the plane and let $\overleftrightarrow{\mathbf{ab}}$ be the line passing through $\mathbf{a}$ and $\mathbf{b}$. The *direction* of the bi-point $[\mathbf{a}, \mathbf{b}]$ is the direction of the line L, that is, the angle $\theta \in [0; \pi[$ that the line $\overleftrightarrow{\mathbf{ab}}$ makes with the positive x-axis (horizontal axis), when measured counterclockwise. The direction of a vector $\overrightarrow{\mathbf{v}} \neq \overrightarrow{\mathbf{0}}$ is the direction of any of its bi-point representatives. See figure 1.4.

5 Definition We say that $[\mathbf{a}, \mathbf{b}]$ has the same direction as $[\mathbf{z}, \mathbf{w}]$ if $\overleftrightarrow{\mathbf{ab}} = \overleftrightarrow{\mathbf{zw}}$. We say that the bi-points $[\mathbf{a}, \mathbf{b}]$ and $[\mathbf{z}, \mathbf{w}]$ have the *same sense* if they have the same direction and if when translating one so as to its tail is over the other's tail, both their heads lie on the same half-plane made by the line perpendicular to then at their tails. They have *opposite sense* if they have the same direction and if when translating one so as to its tail is over the other's tail, their heads lie on different half-planes made by the line perpendicular to them at their tails. . See figures 1.5 and 1.6 . The *sense* of a vector is the sense of any of its bi-point representatives. Two bi-points are *parallel* if the lines containing them are parallel. Two vectors are parallel, if bi-point representatives of them are parallel.

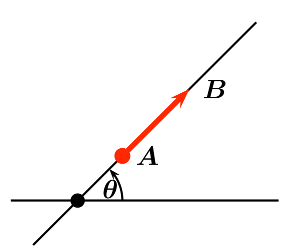

Figure 1.4: Direction of a bi-point

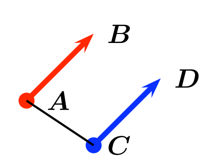

Figure 1.5: Bi-points with the same sense.

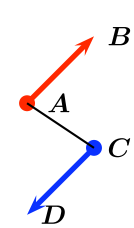

Figure 1.6: Bi-points with opposite sense.

☞ *Bi-point* $[\mathbf{b},\mathbf{a}]$ *has the opposite sense of* $[\mathbf{a},\mathbf{b}]$ *and so we write*

$$[\mathbf{b},\mathbf{a}] = -[\mathbf{a},\mathbf{b}].$$

Similarly we write, $\overrightarrow{\mathbf{ab}} = -\overrightarrow{\mathbf{ba}}$.

6 Definition The *Euclidean length or norm* of bi-point $[\mathbf{a},\mathbf{b}]$ is simply the distance between $\mathbf{a}$ and $\mathbf{b}$ and it is denoted by

$$||[\mathbf{a},\mathbf{b}]|| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}.$$

A bi-point is said to have *unit length* if it has norm 1. The *norm of a vector* is the norm of any of its bi-point representatives.

☞ *A vector is completely determined by three things: (i) its norm, (ii) its direction, and (iii) its sense. It is clear that the norm of a vector satisfies the following properties:*

1. $\left|\left|\overrightarrow{\mathbf{a}}\right|\right| \geq 0$.
2. $\left|\left|\overrightarrow{\mathbf{a}}\right|\right| = 0 \iff \overrightarrow{\mathbf{a}} = \overrightarrow{\mathbf{0}}$.

7 Example The vector $\overrightarrow{\mathbf{v}} = \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$ has norm $\left|\left|\overrightarrow{\mathbf{v}}\right|\right| = \sqrt{1^2 + (\sqrt{2})^2} = \sqrt{3}$.

8 Definition If $\overrightarrow{\mathbf{u}}$ and $\overrightarrow{\mathbf{v}}$ are two vectors in $\mathbb{R}^2$ their *vector sum* $\overrightarrow{\mathbf{u}} + \overrightarrow{\mathbf{v}}$ is defined by the coordinatewise addition

$$\overrightarrow{\mathbf{u}} + \overrightarrow{\mathbf{v}} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}. \tag{1.1}$$

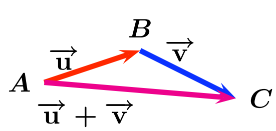

Figure 1.7: Addition of Vectors.

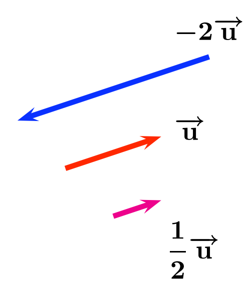

Figure 1.8: Scalar multiplication of vectors.

It is easy to see that vector addition is commutative and associative, that the vector $\overrightarrow{\mathbf{0}}$ acts as an additive identity, and that the additive inverse of $\overrightarrow{\mathbf{a}}$ is $-\overrightarrow{\mathbf{a}}$. To add two vectors geometrically, proceed as follows. Draw a bi-point representative of $\overrightarrow{\mathbf{u}}$. Find a bi-point representative of $\overrightarrow{\mathbf{v}}$ having its tail at the tip of $\overrightarrow{\mathbf{u}}$. The sum $\overrightarrow{\mathbf{u}} + \overrightarrow{\mathbf{v}}$ is the vector whose tail is that of the bi-point for $\overrightarrow{\mathbf{u}}$ and whose tip is that of the bi-point for $\overrightarrow{\mathbf{v}}$. In particular, if $\overrightarrow{\mathbf{u}} = \overrightarrow{\mathbf{AB}}$ and $\overrightarrow{\mathbf{v}} = \overrightarrow{\mathbf{BC}}$, then we have *Chasles' Rule*:

$$\overrightarrow{\mathbf{AB}} + \overrightarrow{\mathbf{BC}} = \overrightarrow{\mathbf{AC}}. \tag{1.2}$$

See figures 1.7, 1.9, 1.10, and 1.11.

9 Definition If $\alpha \in \mathbb{R}$ and $\overrightarrow{\mathbf{a}} \in \mathbb{R}^2$ we define *scalar multiplication* of a vector and a scalar by the coordinatewise multiplication

$$\alpha\overrightarrow{\mathbf{a}} = \alpha \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \alpha a_1 \\ \alpha a_2 \end{bmatrix}. \tag{1.3}$$

It is easy to see that vector addition and scalar multiplication satisfies the following properties.

❶ $\alpha(\overrightarrow{a} + \overrightarrow{b}) = \alpha\overrightarrow{a} + \alpha\overrightarrow{b}$

❷ $(\alpha + \beta)\overrightarrow{a} = \alpha\overrightarrow{a} + \beta\overrightarrow{a}$

❸ $1\overrightarrow{a} = \overrightarrow{a}$

❹ $(\alpha\beta)\overrightarrow{a} = \alpha(\beta\overrightarrow{a})$

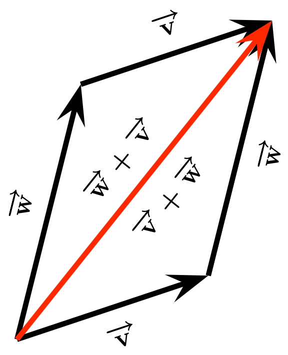

Figure 1.9: Commutativity

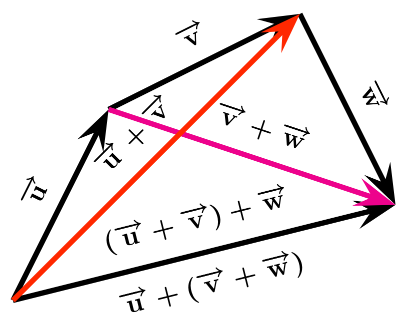

Figure 1.10: Associativity

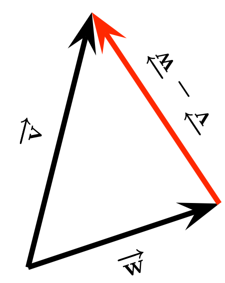

Figure 1.11: Difference

10 Definition Let $\overrightarrow{u} \neq \overrightarrow{0}$. Put $\mathbb{R}\overrightarrow{u} = \{\lambda\overrightarrow{u} : \lambda \in \mathbb{R}\}$ and let $a \in \mathbb{R}^2$, The *affine line with direction vector* $\overrightarrow{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ *and passing through* a is the set of points on the plane

$$a + \mathbb{R}\overrightarrow{u} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x = a_1 + tu_1, \quad y = a_2 + tu_2, \quad t \in \mathbb{R} \right\}.$$

See figure 1.12.

If $u_1 = 0$, the affine line defined above is vertical, as x is constant. If $u_1 \neq 0$, then

$$\frac{x - a_1}{u_1} = t \implies y = a_2 + \frac{(x - a_1)}{u_1}u_2 = \frac{u_2}{u_1}x + a_2 - a_1\frac{u_2}{u_1},$$

that is, the affine line is the Cartesian line with slope $\frac{u_2}{u_1}$. Conversely, if $y = mx + k$ is the equation of a Cartesian line, then

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} 1 \\ m \end{bmatrix} t + \begin{pmatrix} 0 \\ k \end{pmatrix},$$

that is, every Cartesian line is also an affine line and one may take the vector $\begin{bmatrix} 1 \\ m \end{bmatrix}$ as its direction vector. It also follows that two vectors $\overrightarrow{u}$ and $\overrightarrow{v}$ are parallel if and only if the affine lines $\mathbb{R}\overrightarrow{u}$ and $\mathbb{R}\overrightarrow{v}$ are parallel. Hence, $\overrightarrow{u} \parallel \overrightarrow{v}$ if there exists a scalar $\lambda \in \mathbb{R}$ such that $\overrightarrow{u} = \lambda\overrightarrow{v}$.

☞ *Because* $\overrightarrow{0} = 0\overrightarrow{v}$ *for any vector* $\overrightarrow{v}$, *the* $\overrightarrow{0}$ *is parallel to every vector.*

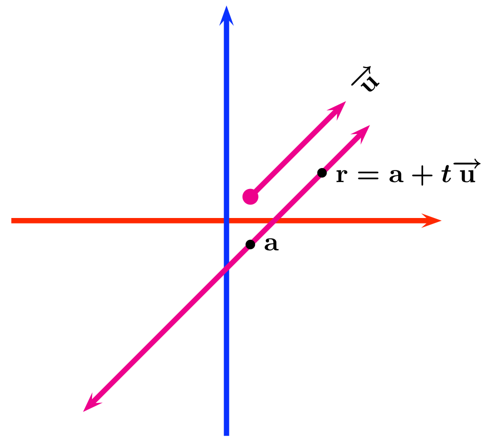

Figure 1.12: Parametric equation of a line on the plane.

11 Example Find a vector of length 3, parallel to $\overrightarrow{\mathbf{v}} = \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$ but in the opposite sense.

Solution: ► *Since* $\left|\left|\overrightarrow{\mathbf{v}}\right|\right| = \sqrt{3}$, *the vector* $\dfrac{\overrightarrow{\mathbf{v}}}{\left|\left|\overrightarrow{\mathbf{v}}\right|\right|}$ *has unit norm and has the same direction and sense as* $\overrightarrow{\mathbf{v}}$, *and so the vector sought is*

$$-3\frac{\overrightarrow{\mathbf{v}}}{\left|\left|\overrightarrow{\mathbf{v}}\right|\right|} = -\frac{3}{\sqrt{3}}\begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} -\sqrt{3} \\ -\sqrt{6} \end{bmatrix}.$$

◄

12 Example Find the parametric equation of the line passing through $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and in the direction of the vector $\begin{bmatrix} 2 \\ -3 \end{bmatrix}$.

Solution: ► *The desired equation is plainly*

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} + t\begin{bmatrix} 2 \\ 3 \end{bmatrix} \implies x = 1 + 2t, \quad y = -1 + 3t, \quad t \in \mathbb{R}.$$

◄

Some plane geometry results can be easily proved by means of vectors. Here are some examples.

13 Example Given a pentagon $ABCDE$, determine the vector sum $\overrightarrow{\mathbf{AB}} + \overrightarrow{\mathbf{BC}} + \overrightarrow{\mathbf{CD}} + \overrightarrow{\mathbf{DE}} + \overrightarrow{\mathbf{EA}}$.

Solution: ► *Utilising Chasles' Rule several times:*

$$\overrightarrow{\mathbf{0}} = \overrightarrow{\mathbf{AA}} = \overrightarrow{\mathbf{AB}} + \overrightarrow{\mathbf{BC}} + \overrightarrow{\mathbf{CD}} + \overrightarrow{\mathbf{DE}} + \overrightarrow{\mathbf{EA}}.$$

. ◄

14 Example Consider a $\triangle \mathbf{ABC}$. Demonstrate that the line segment joining the midpoints of two sides is parallel to the third side and it is in fact, half its length.

Solution: ► *Let the midpoints of* $[\mathbf{A}, \mathbf{B}]$ *and* $[\mathbf{C}, \mathbf{A}]$, *be* $\mathbf{M_C}$ *and* $\mathbf{M_B}$, *respectively. We will demonstrate that* $\overrightarrow{\mathbf{BC}} = 2\overrightarrow{\mathbf{M_C M_B}}$. *We have,* $2\overrightarrow{\mathbf{AM_C}} = \overrightarrow{\mathbf{AB}}$ *and* $2\overrightarrow{\mathbf{AM_B}} = \overrightarrow{\mathbf{AC}}$. *Therefore,*

$$\begin{aligned}
\overrightarrow{\mathbf{BC}} &= \overrightarrow{\mathbf{BA}} + \overrightarrow{\mathbf{AC}} \\
&= -\overrightarrow{\mathbf{AB}} + \overrightarrow{\mathbf{AC}} \\
&= -2\overrightarrow{\mathbf{AM_C}} + 2\overrightarrow{\mathbf{AM_B}} \\
&= 2\overrightarrow{\mathbf{M_C A}} + 2\overrightarrow{\mathbf{AM_B}} \\
&= 2(\overrightarrow{\mathbf{M_C A}} + \overrightarrow{\mathbf{AM_B}}) \\
&= 2\overrightarrow{\mathbf{M_C M_B}},
\end{aligned}$$

as we were to shew. ◄

15 Example In $\triangle ABC$, let $\mathbf{M_C}$ be the midpoint of $[\mathbf{A}, \mathbf{B}]$. Demonstrate that

$$\overrightarrow{\mathbf{CM_C}} = \frac{1}{2}\left(\overrightarrow{\mathbf{CA}} + \overrightarrow{\mathbf{CB}}\right).$$

Solution: ► *As* $\overrightarrow{\mathbf{AM_C}} = \overrightarrow{\mathbf{M_C B}}$, *we have,*

$$\begin{aligned}
\overrightarrow{\mathbf{CA}} + \overrightarrow{\mathbf{CB}} &= \overrightarrow{\mathbf{CM_C}} + \overrightarrow{\mathbf{M_C A}} + \overrightarrow{\mathbf{CM_C}} + \overrightarrow{\mathbf{M_C B}} \\
&= 2\overrightarrow{\mathbf{CM_C}} - \overrightarrow{\mathbf{AM_C}} + \overrightarrow{\mathbf{M_C B}} \\
&= 2\overrightarrow{\mathbf{CM_C}},
\end{aligned}$$

from where the result follows. ◄

16 Example If the medians $[\mathbf{A}, \mathbf{M_A}]$ and $[\mathbf{B}, \mathbf{M_B}]$ of the non-degenerate $\triangle \mathbf{ABC}$ intersect at the point $\mathbf{G}$, demonstrate that

$$\overrightarrow{\mathbf{AG}} = 2\overrightarrow{\mathbf{GM_A}}; \qquad \overrightarrow{\mathbf{BG}} = 2\overrightarrow{\mathbf{GM_B}}.$$

See figure 1.13.

Solution: ► *Since the triangle is non-degenerate, the lines* $\overleftrightarrow{\mathbf{AM_A}}$ *and* $\overleftrightarrow{\mathbf{BM_B}}$ *are not parallel, and hence meet at a point* $\mathbf{G}$. *Therefore,* $\overrightarrow{\mathbf{AG}}$ *and* $\overrightarrow{\mathbf{GM_A}}$ *are parallel and hence there is a scalar* a *such that* $\overrightarrow{\mathbf{AG}} = a\overrightarrow{\mathbf{GM_A}}$. *In the same fashion, there is a scalar* b *such that* $\overrightarrow{\mathbf{BG}} = b\overrightarrow{\mathbf{GM_B}}$. *From example 14,*

$$\begin{aligned}
2\overrightarrow{\mathbf{M_A M_B}} &= \overrightarrow{\mathbf{BA}} \\
&= \overrightarrow{\mathbf{BG}} + \overrightarrow{\mathbf{GA}} \\
&= b\overrightarrow{\mathbf{GM_B}} - a\overrightarrow{\mathbf{GM_A}} \\
&= b\overrightarrow{\mathbf{GM_A}} + b\overrightarrow{\mathbf{M_A M_B}} - a\overrightarrow{\mathbf{GM_A}},
\end{aligned}$$

and thus

$$(2-b)\overrightarrow{\mathbf{M_A M_B}} = (b-a)\overrightarrow{\mathbf{GM_A}}.$$

Since $\triangle \mathbf{ABC}$ *is non-degenerate,* $\overrightarrow{\mathbf{M_A M_B}}$ *and* $\overrightarrow{\mathbf{GM_A}}$ *are not parallel, whence*

$$2-b=0, \quad b-a=0, \quad \implies a=b=2.$$

◄

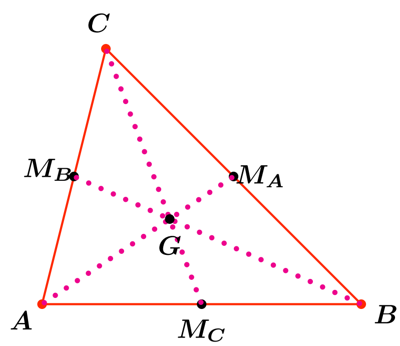

Figure 1.13: Example 17.

17 Example The medians of a non-degenerate triangle $\triangle ABC$ are concurrent. The point of concurrency $\mathbf{G}$ is called the *barycentre* or *centroid* of the triangle. See figure 1.13.

Solution: ▶ *Let* $\mathbf{G}$ *be as in example 16. We must shew that the line* $\overleftrightarrow{\mathbf{CM_C}}$ *also passes through* $\mathbf{G}$*. Let the line* $\overleftrightarrow{\mathbf{CM_C}}$ *and* $\overleftrightarrow{\mathbf{BM_B}}$ *meet in* $\mathbf{G'}$*. By the aforementioned example,*

$$\overrightarrow{\mathbf{AG}} = 2\overrightarrow{\mathbf{GM_A}}; \qquad \overrightarrow{\mathbf{BG}} = 2\overrightarrow{\mathbf{GM_B}}; \quad \overrightarrow{\mathbf{BG'}} = 2\overrightarrow{\mathbf{G'M_B}}; \qquad \overrightarrow{\mathbf{CG'}} = 2\overrightarrow{\mathbf{G'M_C}}.$$

It follows that

$$\begin{aligned}\overrightarrow{\mathbf{GG'}} &= \overrightarrow{\mathbf{GB}} + \overrightarrow{\mathbf{BG'}}\\ &= -2\overrightarrow{\mathbf{GM_B}} + 2\overrightarrow{\mathbf{G'M_B}}\\ &= 2(\overrightarrow{\mathbf{M_BG}} + \overrightarrow{\mathbf{G'M_B}})\\ &= 2\overrightarrow{\mathbf{G'G}}.\end{aligned}$$

Therefore

$$\overrightarrow{\mathbf{GG'}} = -2\overrightarrow{\mathbf{GG'}} \implies 3\overrightarrow{\mathbf{GG'}} = \overrightarrow{0} \implies \overrightarrow{\mathbf{GG'}} = \overrightarrow{0} \implies G = G',$$

demonstrating the result. ◀

Homework

Problem 1.1.1 Is there is any truth to the statement “a vector is that which has magnitude and direction”?

Problem 1.1.2 $ABCD$ is a parallelogram. E is the midpoint of $[\mathbf{B}, \mathbf{C}]$ and F is the midpoint of $[\mathbf{D}, \mathbf{C}]$. Prove that

$$\overrightarrow{\mathbf{AC}} + \overrightarrow{\mathbf{BD}} = 2\overrightarrow{\mathbf{BC}}.$$

Problem 1.1.3 (Varignon's Theorem) Use vector algebra in order to prove that in any quadrilateral $\mathbf{ABCD}$, whose sides do not intersect, the quadrilateral formed by the midpoints of the sides is a parallelogram.

Problem 1.1.4 Let $\mathbf{A}, \mathbf{B}$ be two points on the plane. Construct two points $\mathbf{I}$ and $\mathbf{J}$ such that

$$\overrightarrow{\mathbf{IA}} = -3\overrightarrow{\mathbf{IB}}, \quad \overrightarrow{\mathbf{JA}} = -\frac{1}{3}\overrightarrow{\mathbf{JB}},$$

and then demonstrate that for any arbitrary point $\mathbf{M}$ on the plane

$$\overrightarrow{\mathbf{MA}} + 3\overrightarrow{\mathbf{MB}} = 4\overrightarrow{\mathbf{MI}}$$

and

$$3\overrightarrow{\mathbf{MA}} + \overrightarrow{\mathbf{MB}} = 4\overrightarrow{\mathbf{MJ}}.$$

Problem 1.1.5 Find the Cartesian equation corresponding to the line with parametric equation

$$x = -1 + t, \qquad y = 2 - t.$$

Problem 1.1.6 Let $\mathbf{x}, \mathbf{y}, \mathbf{z}$ be points on the plane with $\mathbf{x} \neq \mathbf{y}$ and consider $\triangle\mathbf{xyz}$. Let $\mathbf{Q}$ be a point on side $[\mathbf{x}, \mathbf{z}]$ such that $||[\mathbf{x}, \mathbf{Q}]|| : ||[\mathbf{Q}, \mathbf{z}]|| = 3 : 4$ and let $\mathbf{P}$ be a point on $[\mathbf{y}, \mathbf{z}]$ such that $||[\mathbf{y}, \mathbf{P}]|| : ||[\mathbf{P}, \mathbf{Q}]|| = 7 : 2$. Let $\mathbf{T}$ be an arbitrary point on the plane.

1. Find rational numbers α and β such that $\overrightarrow{\mathbf{TQ}} = \alpha\overrightarrow{\mathbf{Tx}} + \beta\overrightarrow{\mathbf{Tz}}$.

2. Find rational numbers l, m, n such that $\overrightarrow{\mathbf{TP}} = l\overrightarrow{\mathbf{Tx}} + m\overrightarrow{\mathbf{Ty}} + n\overrightarrow{\mathbf{Tz}}$.

Problem 1.1.7 Let $\mathbf{x}, \mathbf{y}, \mathbf{z}$ be points on the plane with $\mathbf{x} \neq \mathbf{y}$. Demonstrate that

1. The point $\mathbf{a}$ belongs to the line $\overleftrightarrow{\mathbf{xy}}$ if and only if there exists scalars α, β with $\alpha + \beta = 1$ such that
$$\overrightarrow{\mathbf{za}} = \alpha\overrightarrow{\mathbf{zx}} + \beta\overrightarrow{\mathbf{zy}}.$$
2. The point $\mathbf{a}$ belongs to the line segment $[\mathbf{x}; \mathbf{y}]$ if and only if there exists scalars $\alpha \geq 0, \beta \geq 0$ with $\alpha + \beta = 1$ such that
$$\overrightarrow{\mathbf{za}} = \alpha\overrightarrow{\mathbf{zx}} + \beta\overrightarrow{\mathbf{zy}}.$$
3. The point $\mathbf{a}$ belongs to the interior of the triangle $\triangle\mathbf{xyz}$ if and only if there exists scalars $\alpha > 0, \beta > 0$ with $\alpha + \beta < 1$ such that
$$\overrightarrow{\mathbf{za}} = \alpha\overrightarrow{\mathbf{zx}} + \beta\overrightarrow{\mathbf{zy}}.$$

Problem 1.1.8 A circle is divided into three, four equal, or six equal parts (figures 1.17 through 1.19). Find the sum of the vectors. Assume that the divisions start or stop at the centre of the circle, as suggested in the figures.

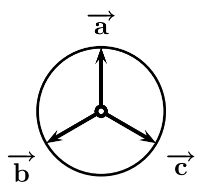

Figure 1.14: [A]. Problem 1.1.8.

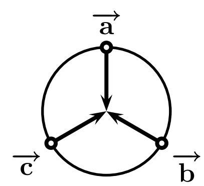

Figure 1.15: [B]. Problem 1.1.8.

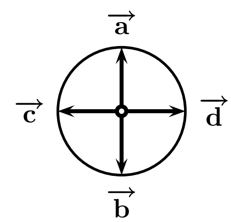

Figure 1.16: [C]. Problem 1.1.8.

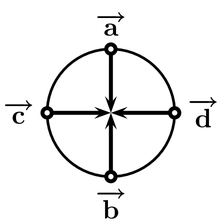

Figure 1.17: [D]. Problem 1.1.8.

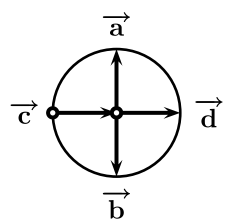

Figure 1.18: [E]. Problem 1.1.8.

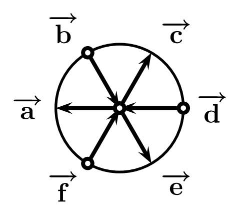

Figure 1.19: [F]. Problem 1.1.8.

1.2 Scalar Product on the Plane

We will now define an operation between two plane vectors that will provide a further tool to examine the geometry on the plane.

18 Definition Let $\overrightarrow{\mathbf{x}} \in \mathbb{R}^2$ and $\overrightarrow{\mathbf{y}} \in \mathbb{R}^2$. Their *scalar product* (*dot product*, *inner product*) is defined and denoted by
$$\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{y}} = x_1y_1 + x_2y_2.$$

19 Example If $\overrightarrow{\mathbf{a}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\overrightarrow{\mathbf{b}} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, then $\overrightarrow{\mathbf{a}} \bullet \overrightarrow{\mathbf{b}} = 1 \cdot 3 + 2 \cdot 4 = 11.$

The following properties of the scalar product are easy to deduce from the definition.

SP1 **Bilinearity**

$$(\overrightarrow{x} + \overrightarrow{y})\bullet\overrightarrow{z} = \overrightarrow{x}\bullet\overrightarrow{z} + \overrightarrow{y}\bullet\overrightarrow{z}, \quad \overrightarrow{x}\bullet(\overrightarrow{y} + \overrightarrow{z}) = \overrightarrow{x}\bullet\overrightarrow{y} + \overrightarrow{x}\bullet\overrightarrow{z} \tag{1.4}$$

SP2 **Scalar Homogeneity**

$$(\alpha\overrightarrow{x})\bullet\overrightarrow{y} = \overrightarrow{x}\bullet(\alpha\overrightarrow{y}) = \alpha(\overrightarrow{x}\bullet\overrightarrow{y}),\ \alpha \in \mathbb{R}. \tag{1.5}$$

SP3 **Commutativity**

$$\overrightarrow{x}\bullet\overrightarrow{y} = \overrightarrow{y}\bullet\overrightarrow{x} \tag{1.6}$$

SP4

$$\overrightarrow{x}\bullet\overrightarrow{x} \geq 0 \tag{1.7}$$

SP5

$$\overrightarrow{x}\bullet\overrightarrow{x} = 0 \iff \overrightarrow{x} = \overrightarrow{0} \tag{1.8}$$

SP6

$$||\overrightarrow{x}|| = \sqrt{\overrightarrow{x}\bullet\overrightarrow{x}} \tag{1.9}$$

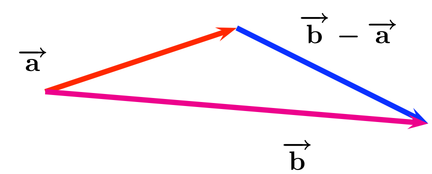

Figure 1.20: Theorem 21.

20 Definition Given vectors $\overrightarrow{a}$ and $\overrightarrow{b}$, we define the (convex) angle between them, denoted by $(\widehat{\overrightarrow{a}, \overrightarrow{b}}) \in [0; \pi]$, as the angle between the affine lines $\mathbb{R}\overrightarrow{a}$ and $\mathbb{R}\overrightarrow{b}$.

21 Theorem Let $\overrightarrow{a}$ and $\overrightarrow{b}$ be vectors in $\mathbb{R}^2$. Then

$$\overrightarrow{a}\bullet\overrightarrow{b} = ||\overrightarrow{a}||||\overrightarrow{b}||\cos(\widehat{\overrightarrow{a}, \overrightarrow{b}}).$$

Proof: *From figure 1.20, using Al-Kashi's Law of Cosines on the length of the vectors, and (1.4) through (1.9) we have*

$$\begin{aligned}
& ||\overrightarrow{b} - \overrightarrow{a}||^2 = ||\overrightarrow{a}||^2 + ||\overrightarrow{b}||^2 - 2||\overrightarrow{a}||||\overrightarrow{b}||\cos(\widehat{\overrightarrow{a}, \overrightarrow{b}}) \\
\iff & (\overrightarrow{b} - \overrightarrow{a})\bullet(\overrightarrow{b} - \overrightarrow{a}) = ||\overrightarrow{a}||^2 + ||\overrightarrow{b}||^2 - 2||\overrightarrow{a}||||\overrightarrow{b}||\cos(\widehat{\overrightarrow{a}, \overrightarrow{b}}) \\
\iff & \overrightarrow{b}\bullet\overrightarrow{b} - 2\overrightarrow{a}\bullet\overrightarrow{b} + \overrightarrow{a}\bullet\overrightarrow{a} = ||\overrightarrow{a}||^2 + ||\overrightarrow{b}||^2 - 2||\overrightarrow{a}||||\overrightarrow{b}||\cos(\widehat{\overrightarrow{a}, \overrightarrow{b}}) \\
\iff & ||\overrightarrow{b}||^2 - 2\overrightarrow{a}\bullet\overrightarrow{b} + ||\overrightarrow{a}||^2 = ||\overrightarrow{a}||^2 + ||\overrightarrow{b}||^2 - 2||\overrightarrow{a}||||\overrightarrow{b}||\cos(\widehat{\overrightarrow{a}, \overrightarrow{b}}) \\
\iff & \overrightarrow{a}\bullet\overrightarrow{b} = ||\overrightarrow{a}||||\overrightarrow{b}||\cos(\widehat{\overrightarrow{a}, \overrightarrow{b}}),
\end{aligned}$$

as we wanted to shew. ❑

Putting $(\widehat{\overrightarrow{a}, \overrightarrow{b}}) = \frac{\pi}{2}$ in Theorem 21 we obtain the following corollary.

22 Corollary Two vectors in $\mathbb{R}^2$ are perpendicular if and only if their dot product is 0.

☞ *It follows that the vector $\overrightarrow{0}$ is simultaneously parallel and perpendicular to any vector!*

23 Definition Two vectors are said to be *orthogonal* if they are perpendicular. If $\overrightarrow{a}$ is orthogonal to $\overrightarrow{b}$, we write $\overrightarrow{a} \perp \overrightarrow{b}$.

24 Definition If $\overrightarrow{a} \perp \overrightarrow{b}$ and $||\overrightarrow{a}|| = \left|\left|\overrightarrow{b}\right|\right| = 1$ we say that $\overrightarrow{a}$ and $\overrightarrow{b}$ are *orthonormal.*

Since $|\cos\theta| \leq 1$ we also have

25 Corollary (Cauchy-Bunyakovsky-Schwarz Inequality)

$$\left|\overrightarrow{a}\bullet\overrightarrow{b}\right| \leq ||\overrightarrow{a}||\left|\left|\overrightarrow{b}\right|\right|.$$

Equality occurs if and only if $\overrightarrow{a} \parallel \overrightarrow{b}$.

If $\overrightarrow{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\overrightarrow{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, the CBS Inequality takes the form

$$|a_1b_1 + a_2b_2| \leq (a_1^2 + a_2^2)^{1/2}(b_1^2 + b_2^2)^{1/2}. \tag{1.10}$$

26 Example Let a, b be positive real numbers. Minimise $a^2 + b^2$ subject to the constraint $a + b = 1$.

Solution: ▶ *By the CBS Inequality,*

$$1 = |a \cdot 1 + b \cdot 1| \leq (a^2 + b^2)^{1/2}(1^2 + 1^2)^{1/2} \implies a^2 + b^2 \geq \frac{1}{2}.$$

Equality occurs if and only if $\begin{bmatrix} a \\ b \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. *In such case,* $a = b = \lambda$, *and so equality is achieved for* $a = b = \frac{1}{2}$. ◀

27 Corollary (Triangle Inequality)

$$\left|\left|\overrightarrow{a} + \overrightarrow{b}\right|\right| \leq ||\overrightarrow{a}|| + \left|\left|\overrightarrow{b}\right|\right|.$$

Proof:

$$\begin{aligned} ||\overrightarrow{a} + \overrightarrow{b}||^2 &= (\overrightarrow{a} + \overrightarrow{b})\bullet(\overrightarrow{a} + \overrightarrow{b}) \\ &= \overrightarrow{a}\bullet\overrightarrow{a} + 2\,\overrightarrow{a}\bullet\overrightarrow{b} + \overrightarrow{b}\bullet\overrightarrow{b} \\ &\leq ||\overrightarrow{a}||^2 + 2||\overrightarrow{a}||||\overrightarrow{b}|| + ||\overrightarrow{b}||^2 \\ &= (||\overrightarrow{a}|| + ||\overrightarrow{b}||)^2, \end{aligned}$$

from where the desired result follows. ❑

28 Example Let x, y, z be positive real numbers. Prove that

$$\sqrt{2}(x + y + z) \leq \sqrt{x^2 + y^2} + \sqrt{y^2 + z^2} + \sqrt{z^2 + x^2}.$$

Solution: ▶ *Put* $\overrightarrow{a} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\overrightarrow{b} = \begin{bmatrix} y \\ z \end{bmatrix}$, $\overrightarrow{c} = \begin{bmatrix} z \\ x \end{bmatrix}$. *Then*

$$\left|\left|\overrightarrow{a} + \overrightarrow{b} + \overrightarrow{c}\right|\right| = \left|\left|\begin{bmatrix} x + y + z \\ x + y + z \end{bmatrix}\right|\right| = \sqrt{2}(x + y + z).$$

Also,

$$\left|\left|\overrightarrow{\mathbf{a}}\right|\right| + \left|\left|\overrightarrow{\mathbf{b}}\right|\right| + \left|\left|\overrightarrow{\mathbf{c}}\right|\right| = \sqrt{x^2+y^2} + \sqrt{y^2+z^2} + \sqrt{z^2+x^2},$$

and the assertion follows by the triangle inequality

$$\left|\left|\overrightarrow{\mathbf{a}} + \overrightarrow{\mathbf{b}} + \overrightarrow{\mathbf{c}}\right|\right| \leq \left|\left|\overrightarrow{\mathbf{a}}\right|\right| + \left|\left|\overrightarrow{\mathbf{b}}\right|\right| + \left|\left|\overrightarrow{\mathbf{c}}\right|\right|.$$

◀

We now use vectors to prove a classical theorem of Euclidean geometry.

29 Definition Let A and B be points on the plane and let $\overrightarrow{\mathbf{u}}$ be a unit vector. If $\overrightarrow{\mathbf{AB}} = \lambda\overrightarrow{\mathbf{u}}$, then λ is the *directed distance or algebraic measure of the line segment* $[AB]$ *with respect to the vector* $\overrightarrow{\mathbf{u}}$. We will denote this distance by $\overline{AB}_{\overrightarrow{\mathbf{u}}}$, or more routinely, if the vector $\overrightarrow{\mathbf{u}}$ is patent, by $\overline{AB}$. Observe that $\overline{AB} = -\overline{BA}$.

30 Theorem (Thales' Theorem) Let $\overleftrightarrow{\mathbf{D}}$ y $\overleftrightarrow{\mathbf{D'}}$ be two distinct lines on the plane. Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$ be distinct points of $\overleftrightarrow{\mathbf{D}}$, and $\mathbf{A'}, \mathbf{B'}, \mathbf{C'}$ be distinct points of $\overleftrightarrow{\mathbf{D'}}$, $\mathbf{A} \neq \mathbf{A'}$, $\mathbf{B} \neq \mathbf{B'}$, $\mathbf{C} \neq \mathbf{C'}$, $\mathbf{A} \neq \mathbf{B}$, $\mathbf{A'} \neq \mathbf{B'}$. Let $\overleftrightarrow{\mathbf{AA'}} \parallel \overleftrightarrow{\mathbf{BB'}}$. Then

$$\overleftrightarrow{\mathbf{AA'}} \parallel \overleftrightarrow{\mathbf{CC'}} \iff \frac{\overline{AC}}{\overline{AB}} = \frac{\overline{A'C'}}{\overline{A'B'}}.$$

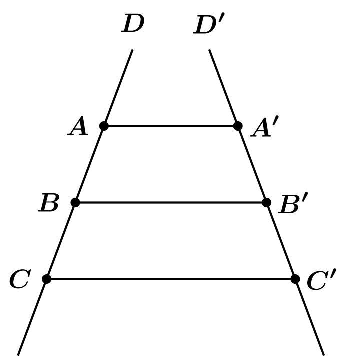

Figure 1.21: Thales' Theorem.

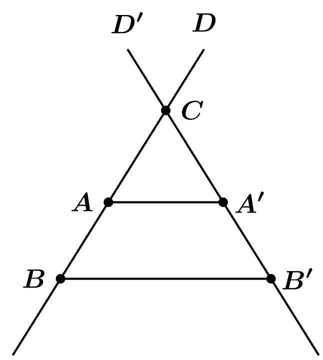

Figure 1.22: Corollary to Thales' Theorem.

Proof: *Refer to figure 1.2. On the one hand, because they are unit vectors in the same direction,*

$$\frac{\overrightarrow{\mathbf{AB}}}{\overline{AB}} = \frac{\overrightarrow{\mathbf{AC}}}{\overline{AC}}; \qquad \frac{\overrightarrow{\mathbf{A'B'}}}{\overline{A'B'}} = \frac{\overrightarrow{\mathbf{A'C'}}}{\overline{A'C'}}.$$

On the other hand, by Chasles' Rule,

$$\overrightarrow{\mathbf{BB'}} = \overrightarrow{\mathbf{BA}} + \overrightarrow{\mathbf{AA'}} + \overrightarrow{\mathbf{A'B'}} = (\overrightarrow{\mathbf{A'B'}} - \overrightarrow{\mathbf{AB}}) + \overrightarrow{\mathbf{AA'}}.$$

Since $\overleftrightarrow{\mathbf{AA'}} \parallel \overleftrightarrow{\mathbf{BB'}}$, *there is a scalar* $\lambda \in \mathbb{R}$ *such that*

$$\overrightarrow{\mathbf{A'B'}} = \overrightarrow{\mathbf{AB}} + \lambda\overrightarrow{\mathbf{AA'}}.$$

Assembling these results,

$$\begin{aligned}
\overrightarrow{\mathbf{CC'}} &= \overrightarrow{\mathbf{CA}} + \overrightarrow{\mathbf{AA'}} + \overrightarrow{\mathbf{A'C'}} \\
&= -\frac{\overline{AC}}{\overline{AB}} \cdot \overrightarrow{\mathbf{AB}} + \overrightarrow{\mathbf{AA'}} + \frac{\overline{A'C'}}{\overline{A'B'}}\left(\overrightarrow{\mathbf{AB}} + \lambda\overrightarrow{\mathbf{AA'}}\right) \\
&= \left(\frac{\overline{A'C'}}{\overline{A'B'}} - \frac{\overline{AC}}{\overline{AB}}\right)\overrightarrow{\mathbf{AB}} + \left(1 + \lambda\frac{\overline{A'C'}}{\overline{A'B'}}\right)\overrightarrow{\mathbf{AA'}}.
\end{aligned}$$

As the line $\overleftrightarrow{AA'}$ is not parallel to the line $\overleftrightarrow{AB}$, the equality above reveals that

$$\overleftrightarrow{AA'} \parallel \overleftrightarrow{CC'} \iff \frac{\overline{AC}}{\overline{AB}} - \frac{\overline{A'C'}}{\overline{A'B'}} = 0,$$

proving the theorem.❑

From the preceding theorem, we immediately gather the following corollary. (See figure 1.2.)

31 Corollary Let $\overleftrightarrow{D}$ and $\overleftrightarrow{D'}$ are distinct lines, intersecting in the unique point C. Let A, B, be points on line $\overleftrightarrow{D}$, and A', B', points on line $\overleftrightarrow{D'}$. Then

$$\overleftrightarrow{AA'} \parallel \overleftrightarrow{BB'} \iff \frac{\overline{CB}}{\overline{CA}} = \frac{\overline{CB'}}{\overline{CA'}}.$$

1.3 Linear Independence

Consider now two arbitrary vectors in $\mathbb{R}^2$, $\overrightarrow{x}$ and $\overrightarrow{y}$, say. Under which conditions can we write an arbitrary vector $\overrightarrow{v}$ on the plane as a linear combination of $\overrightarrow{x}$ and $\overrightarrow{y}$, that is, when can we find scalars a, b such that

$$\overrightarrow{v} = a\overrightarrow{x} + b\overrightarrow{y}\ ?$$

The answer can be promptly obtained algebraically. Operating formally,

$$\begin{aligned}\overrightarrow{v} = a\overrightarrow{x} + b\overrightarrow{y} \quad &\iff \quad v_1 = ax_1 + by_1, \qquad v_2 = ax_2 + by_2 \\ &\iff \quad a = \frac{v_1y_2 - v_2y_1}{x_1y_2 - x_2y_1}, \qquad b = \frac{x_1v_2 - x_2v_1}{x_1y_2 - x_2y_1}.\end{aligned}$$

The above expressions for a and b make sense only if $x_1y_2 \neq x_2y_1$. But, what does it mean $x_1y_2 = x_2y_1$? If none of these are zero then $\dfrac{x_1}{y_1} = \dfrac{x_2}{y_2} = \lambda$, say, and to

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \iff \overrightarrow{x} \parallel \overrightarrow{y}.$$

If $x_1 = 0$, then either $x_2 = 0$ or $y_1 = 0$. In the first case, $\overrightarrow{x} = \overrightarrow{0}$, and *a fortiori* $\overrightarrow{x} \parallel \overrightarrow{y}$, since all vectors are parallel to the zero vector. In the second case we have

$$\overrightarrow{x} = x_2\overrightarrow{j}, \qquad \overrightarrow{y} = y_2\overrightarrow{j},$$

and so both vectors are parallel to $\overrightarrow{j}$ and hence $\overrightarrow{x} \parallel \overrightarrow{y}$. We have demonstrated the following theorem.

32 Theorem Given two vectors in $\mathbb{R}^2$, $\overrightarrow{x}$ and $\overrightarrow{y}$, an arbitrary vector $\overrightarrow{v}$ can be written as the linear combination

$$\overrightarrow{v} = a\overrightarrow{x} + b\overrightarrow{y}, \quad a \in \mathbb{R}, \quad b \in \mathbb{R}$$

if and only if $\overrightarrow{x}$ is not parallel to $\overrightarrow{y}$. In this last case we say that $\overrightarrow{x}$ is *linearly independent* from vector $\overrightarrow{y}$. If two vectors are not linearly independent, then we say that they are *linearly dependent.*

33 Example The vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ are clearly linearly independent, since one is not a scalar multiple of the other. Given an arbitrary vector $\begin{bmatrix} a \\ b \end{bmatrix}$ we can express it as a linear combination of these vectors as follows:

$$\begin{bmatrix} a \\ b \end{bmatrix} = (a-b)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Consider now two linearly independent vectors $\overrightarrow{x}$ and $\overrightarrow{y}$. For $a \in [0;1]$, $a\overrightarrow{x}$ is parallel to $\overrightarrow{x}$ and traverses the whole length of $\overrightarrow{x}$: from its tip (when $a = 1$) to its tail (when $a = 0$). In the same manner, for $b \in [0;1]$, $b\overrightarrow{y}$ is parallel to $\overrightarrow{y}$ and traverses the whole length of $\overrightarrow{y}$. The linear combination $a\overrightarrow{x} + b\overrightarrow{y}$ is also a vector on the plane.

34 Definition Given two linearly independent vectors $\overrightarrow{x}$ and $\overrightarrow{y}$ consider bi-point representatives of them with the tails at the origin. The *fundamental parallelogram* of the the vectors $\overrightarrow{x}$ and $\overrightarrow{y}$ is the set

$$\{a\overrightarrow{x} + b\overrightarrow{y} : a \in [0;1], b \in [0;1]\}.$$

Figure 1.23 shews the fundamental parallelogram of $\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\}$, coloured in brown, and the respective tiling of the plane by various translations of it. Observe that the vertices of this parallelogram are $\left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\1\end{pmatrix}\right\}$. In essence then, linear independence of two vectors on the plane means that we may obtain every vector on the plane as a linear combination of these two vectors and hence cover the whole plane by all these linear combinations.

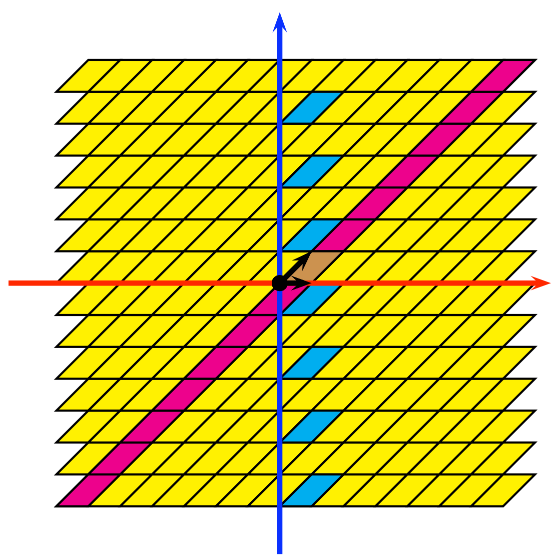

Figure 1.23: Tiling and the fundamental parallelogram.

Homework

Problem 1.3.1 Prove that $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}-1\\1\end{bmatrix}$ are linearly independent, and draw their fundamental parallelogram.

Problem 1.3.2 Write an arbitrary vector $\begin{bmatrix}a\\b\end{bmatrix}$ on the plane, as a linear combination of the vectors $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}-1\\1\end{bmatrix}$.

Problem 1.3.3 Consider the line with Cartesian equation $L : ax + by = c$, where not both of a, b are zero. Let $\mathbf{t}$ be a point not on L. Find a formula for the distance from $\mathbf{t}$ to L.

Problem 1.3.4 Prove that two non-zero perpendicular vectors in $\mathbb{R}^2$ must be linearly independent.

1.4 Geometric Transformations in two dimensions

We now are interested in the following fundamental functions of sets (figures) on the plane: translations, scalings (stretching or shrinking) reflexions about the axes, and rotations about the origin. It will turn out that a handy tool for investigating all of these (with the exception of translations), will be certain construct called *matrices* which we will study in the next section.

First observe what is meant by a function $F : \mathbb{R}^2 \to \mathbb{R}^2$. This means that the input of the function is a point of the plane, and the output is also a point on the plane.

A rather uninteresting example, but nevertheless an important one is the following.

35 Example The function $\mathbf{I} : \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{I}(\mathbf{x}) = \mathbf{x}$ is called the *identity transformation*. Observe that the identity transformation leaves a point untouched.

We start with the simplest of these functions.

36 Definition A function $T_{\overrightarrow{v}} : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be a *translation* if it is of the form $T_{\overrightarrow{v}}(\mathbf{x}) = \mathbf{x} + \overrightarrow{v}$, where $\overrightarrow{v}$ is a fixed vector on the plane.

A translation simply shifts an object on the plane rigidly (that is, it does not distort it shape or re-orient it), to a copy of itself a given amount of units from where it was. See figure 1.24 for an example.

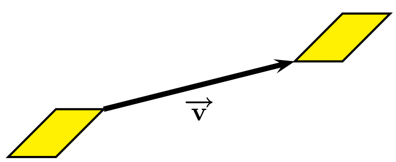

Figure 1.24: A translation.

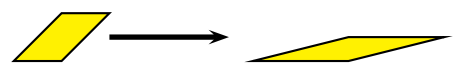

Figure 1.25: A scaling.

It is clear that the composition of any two translations commutes, that is, if $T_{\overrightarrow{v}_1}, T_{\overrightarrow{v}_2} : \mathbb{R}^2 \to \mathbb{R}^2$ are translations, then $T_{\overrightarrow{v}_1} \circ T_{\overrightarrow{v}_2} = T_{\overrightarrow{v}_2} \circ T_{\overrightarrow{v}_1}$. For let $T_1(\mathbf{a}) = \mathbf{a} + \overrightarrow{v}_1$ and $T_{\overrightarrow{v}_2}(\mathbf{a}) = \mathbf{a} + \overrightarrow{v}_2$. Then

$$(T_{\overrightarrow{v}_1} \circ T_{\overrightarrow{v}_2})(\mathbf{a}) = T_{\overrightarrow{v}_1}(T_{\overrightarrow{v}_2}(\mathbf{a})) = T_{\overrightarrow{v}_2}(\mathbf{a}) + \overrightarrow{v}_1 = \mathbf{a} + \overrightarrow{v}_2 + \overrightarrow{v}_1,$$

and

$$(T_{\overrightarrow{v}_2} \circ T_{\overrightarrow{v}_1})(\mathbf{a}) = T_{\overrightarrow{v}_2}(T_{\overrightarrow{v}_1}(\mathbf{a})) = T_{\overrightarrow{v}_1}(\mathbf{a}) + \overrightarrow{v}_2 = \mathbf{a} + \overrightarrow{v}_1 + \overrightarrow{v}_2,$$

from where the commutativity claim is deduced.

37 Definition A function $S_{a,b} : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be a *scaling* if it is of the form $S_{a,b}(\mathbf{r}) = \begin{pmatrix} ax \\ by \end{pmatrix}$, where $a > 0$, $b > 0$ are real numbers.

Figure 1.25 shews the scaling $S_{2,0.5}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} 2x \\ 0.5y \end{pmatrix}$.

It is clear that the composition of any two scalings commutes, that is, if $S_{a,b}, S_{a',b'} : \mathbb{R}^2 \to \mathbb{R}^2$ are scalings, then $S_{a,b} \circ S_{a',b'} = S_{a',b'} \circ S_{a,b}$. For

$$(S_{a,b} \circ S_{a',b'})(\mathbf{r}) = S_{a,b}(S_{a',b'}(\mathbf{r})) = S_{a,b}\left(\begin{pmatrix} a'x \\ b'y \end{pmatrix}\right) = \begin{pmatrix} a(a'x) \\ b(b'y) \end{pmatrix},$$

and

$$(S_{a',b'} \circ S_{a,b})(\mathbf{r}) = S_{a',b'}(S_{a,b}(\mathbf{r})) = S_{a',b'}\left(\begin{pmatrix} ax \\ by \end{pmatrix}\right) = \begin{pmatrix} a'(ax) \\ b'(by) \end{pmatrix},$$

from where the commutativity claim is deduced.

Translations and scalings do not necessarily commute, however. For consider the translation $T_{\overrightarrow{\mathrm{i}}}(\mathbf{a}) = \mathbf{a} + \overrightarrow{\mathrm{i}}$ and the scaling $S_{2,1}(\mathbf{a}) = \begin{pmatrix} 2a_1 \\ a_2 \end{pmatrix}$. Then

$$(T_{\overrightarrow{\mathrm{i}}} \circ S_{2,1})\left(\begin{pmatrix} -1 \\ 0 \end{pmatrix}\right) = T_{\overrightarrow{\mathrm{i}}}\left(S\left(\begin{pmatrix} -1 \\ 0 \end{pmatrix}\right)\right) = T_{\overrightarrow{\mathrm{i}}}\left(\begin{pmatrix} -2 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

but

$$(S_{2,1} \circ T_{\overrightarrow{\mathrm{i}}})\left(\begin{pmatrix} -1 \\ 0 \end{pmatrix}\right) = S_{2,1}\left(T_{\overrightarrow{\mathrm{i}}}\left(\begin{pmatrix} -1 \\ 0 \end{pmatrix}\right)\right) = S_{2,1}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

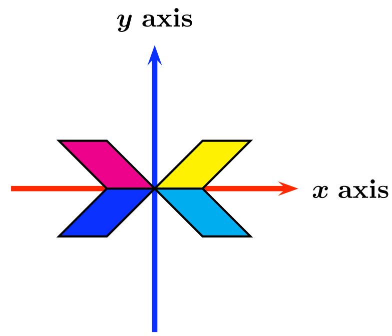

Figure 1.26: Reflexions. The original object (in the first quadrant) is yellow. Its reflexion about the y-axis is magenta (on the second quadrant). Its reflexion about the x-axis is cyan (on the fourth quadrant). Its reflexion about the origin is blue (on the third quadrant).

38 Definition A function $R_H : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be a *reflexion about the y-axis or horizontal reflexion* if it is of the form $R_H(\mathbf{r}) = \begin{pmatrix} -x \\ y \end{pmatrix}$. A function $R_V : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be a *reflexion about the x-axis or vertical reflexion* if it is of the form $R_V(\mathbf{r}) = \begin{pmatrix} x \\ -y \end{pmatrix}$. A function $R_O : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be a *reflexion about origin* if it is of the form $R_H(\mathbf{r}) = \begin{pmatrix} -x \\ -y \end{pmatrix}$.

Some reflexions appear in figure 1.26.

A few short computations establish various commutativity properties among reflexions, translations, and scalings. See problem 1.4.4.

We now define rotations. This definition will be somewhat harder than the others, so let us develop some ancillary results.

Consider a point $\mathbf{r}$ with polar coordinates $x = \rho\cos\alpha$, $y = \rho\sin\alpha$ as in figure 1.27. Here $\rho = \sqrt{x^2 + y^2}$ and $\alpha \in [0; 2\pi[$. If we rotate it, in the levogyrate sense, by an angle θ, we land on the new

point $\mathbf{x}'$ with $x' = \rho\cos(\alpha+\theta)$ and $y' = \rho\sin(\alpha+\theta)$. But

$$\rho\cos(\alpha+\theta) = \rho\cos\theta\cos\alpha - \rho\sin\theta\sin\alpha = x\cos\theta - y\sin\theta,$$

and

$$\rho\sin(\alpha+\theta) = \rho\sin\alpha\cos\theta + \rho\sin\theta\cos\alpha = y\cos\theta + x\sin\theta.$$

Hence the point $\begin{pmatrix} x \\ y \end{pmatrix}$ is mapped to the point $\begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}$.

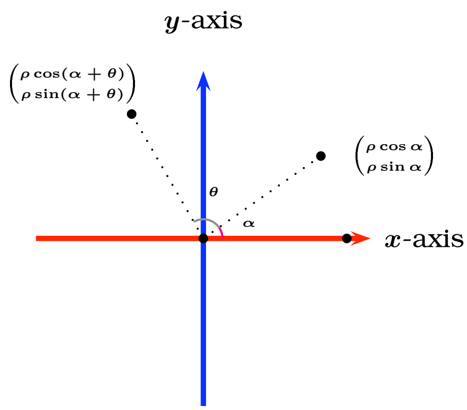

Figure 1.27: Rotation by an angle θ in the levogyrate (counterclockwise) sense from the x-axis.

We may now formulate the definition of a rotation.

39 Definition A function $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be a *levogyrate rotation about the origin by the angle θ measured from the positive x-axis* if $R_\theta(\mathbf{r}) = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}$. Here $\rho = \sqrt{x^2+y^2}$.

Various properties of the composition of rotations with other plane transformations are explored in problems 1.4.5 and 1.4.6.

We now codify some properties shared by scalings, reflexions, and rotations.

40 Definition A function $L : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be a *linear transformation* from $\mathbb{R}^2$ to $\mathbb{R}^2$ if for all points $\mathbf{a}, \mathbf{b}$ on the plane and every scalar λ, it is verified that

$$L(\mathbf{a}+\mathbf{b}) = L(\mathbf{a}) + L(\mathbf{b}), \qquad L(\lambda\mathbf{a}) = \lambda L(\mathbf{a}).$$

It is easy to prove that scalings, reflexions, and rotations are linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$, but not so translations.

41 Definition A function $A : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be an *affine transformation* from $\mathbb{R}^2$ to $\mathbb{R}^2$ if there exists a linear transformation $L : \mathbb{R}^2 \to \mathbb{R}^2$ and a fixed vector $\overrightarrow{\mathbf{v}} \in \mathbb{R}^2$ such that for all points $\mathbf{x} \in \mathbb{R}^2$ it is verified that

$$A(\mathbf{x}) = L(\mathbf{x}) + \overrightarrow{\mathbf{v}}.$$

It is easy to see that translations are then affine transformations, where for the linear transformation L in the definition we may take $I : \mathbb{R}^2 \to \mathbb{R}^2$, the identity transformation $I(\mathbf{x}) = \mathbf{x}$.

We have seen that scalings, reflexions and rotations are linear transformations. If $L : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation, then

$$L(\mathbf{r}) = L(x\overrightarrow{\mathbf{i}} + y\overrightarrow{\mathbf{j}}) = xL(\overrightarrow{\mathbf{i}}) + yL(\overrightarrow{\mathbf{j}}),$$

and thus a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ is solely determined by the values $L(\overrightarrow{i})$ and $L(\overrightarrow{j})$. We will now introduce a way to codify these values.

42 Definition Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation. The *matrix A_L associated to L* is the 2×2, (2 rows, 2 columns) array whose columns are (in this order) $L\left(\begin{pmatrix}1\\0\end{pmatrix}\right)$ and $L\left(\begin{pmatrix}0\\1\end{pmatrix}\right)$.

43 Example (Scaling Matrices) Let $a > 0, b > 0$ be a real numbers. The matrix of the scaling transformation $S_{a,b}$ is $\begin{bmatrix}a & 0\\0 & b\end{bmatrix}$. For

$$S_{a,b}\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}a \cdot 1\\b \cdot 0\end{pmatrix} = \begin{pmatrix}a\\0\end{pmatrix}$$

and

$$S_{a,b}\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}0 \cdot 1\\b \cdot 1\end{pmatrix} = \begin{pmatrix}0\\b\end{pmatrix}.$$

44 Example (Reflexion Matrices) It is easy to verify that the matrix for the transformation R_H is $\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}$, that the matrix for the transformation R_V is $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$, and the matrix for the transformation R_O is $\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$.

45 Example (Rotating Matrices) It is easy to verify that the matrix for a rotation R_θ is $\begin{bmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{bmatrix}$.

46 Example (Identity Matrix) The matrix for the identity linear transformation $\mathbf{Id} : \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{Id}(\mathbf{x}) = \mathbf{x}$ is $\mathbf{I}_2 = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$.

47 Example (Zero Matrix) The matrix for the null linear transformation $\mathbf{N} : \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{N}(\mathbf{x}) = \mathbf{O}$ is $\mathbf{0}_2 = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$.

From problem 1.4.7 we know that the composition of two linear transformations is also linear. We are now interested in how to codify the matrix of a composition of linear transformations $L_1 \circ L_1$ in terms of their individual matrices.

48 Theorem Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ have the matrix representation $A_L = \begin{bmatrix}a & b\\c & d\end{bmatrix}$ and let $L' : \mathbb{R}^2 \to \mathbb{R}^2$ have the matrix representation $A_{L'} = \begin{bmatrix}r & s\\t & u\end{bmatrix}$. Then the composition $L \circ L'$ has matrix representation

$$\begin{bmatrix}ar + bt & as + bu\\cr + dt & cs + du\end{bmatrix}.$$

Proof: *We need to find* $(L \circ L')\left(\begin{pmatrix}1\\0\end{pmatrix}\right)$ *and* $(L \circ L')\left(\begin{pmatrix}0\\1\end{pmatrix}\right)$.

We have

$$(L\circ L')\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = L\left(L'\left(\begin{pmatrix}1\\0\end{pmatrix}\right)\right) = L\left(\begin{pmatrix}r\\t\end{pmatrix}\right) = rL(\overrightarrow{\mathrm{i}})+tL(\overrightarrow{\mathrm{j}}) = r\begin{pmatrix}a\\c\end{pmatrix}+t\begin{pmatrix}b\\d\end{pmatrix} = \begin{pmatrix}ar+bt\\cr+dt\end{pmatrix},$$

and

$$(L\circ L')\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = L\left(L'\left(\begin{pmatrix}0\\1\end{pmatrix}\right)\right) = L\left(\begin{pmatrix}s\\u\end{pmatrix}\right) = sL(\overrightarrow{\mathrm{i}})+uL(\overrightarrow{\mathrm{j}}) = s\begin{pmatrix}a\\c\end{pmatrix}+u\begin{pmatrix}b\\d\end{pmatrix} = \begin{pmatrix}as+bu\\cs+du\end{pmatrix},$$

whence we conclude that the matrix of $L \circ L'$ *is* $\begin{bmatrix}ar+bt & as+bu\\ cr+dt & cs+du\end{bmatrix}$, *as we wanted to shew.*❑

The above motivates the following definition.

49 Definition Let $A = \begin{bmatrix}a & b\\ c & d\end{bmatrix}$ and $B = \begin{bmatrix}r & s\\ t & u\end{bmatrix}$ be two 2×2 matrices, and $\lambda \in \mathbb{R}$ be a scalar. We define *matrix addition* as

$$A+B = \begin{bmatrix}a & b\\ c & d\end{bmatrix} + \begin{bmatrix}r & s\\ t & u\end{bmatrix} = \begin{bmatrix}a+r & b+s\\ c+t & d+u\end{bmatrix}.$$

We define *matrix multiplication* as

$$AB = \begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}r & s\\ t & u\end{bmatrix} = \begin{bmatrix}ar+bt & as+bu\\ cr+dt & cs+du\end{bmatrix}.$$

We define *scalar multiplication of a matrix* as

$$\lambda A = \lambda\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}\lambda a & \lambda b\\ \lambda c & \lambda d\end{bmatrix}.$$

☞ *Since the composition of functions is not necessarily commutative, neither is matrix multiplication. Since the composition of functions is associative, so is matrix multiplication.*

50 Example Let

$$M = \begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix}, \qquad N = \begin{bmatrix}1 & 2\\ -2 & 1\end{bmatrix}.$$

Then

$$M+N = \begin{bmatrix}2 & 1\\ -2 & 2\end{bmatrix}, \qquad 2M = \begin{bmatrix}2 & -2\\ 0 & 2\end{bmatrix}, \qquad MN = \begin{bmatrix}3 & 1\\ -2 & 1\end{bmatrix}.$$

51 Example Find a 2×2 matrix that will transform the square in figure 1.28 into the parallelogram in figure 1.29. Assume in each case that the vertices of the figures are lattice points, that is, coordinate points with integer coordinates.

Solution: ► *Let* $\begin{bmatrix}a & b\\ c & d\end{bmatrix}$ *be the desired matrix. Then since*

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix},$$

the point $\begin{pmatrix}0\\0\end{pmatrix}$ *is a* fortiori, *transformed to itself. We now assume, without loss of generality, that each vertex of the square is transformed in the same order, counterclockwise, to each vertex of the rectangle. Then*

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}2\\2\end{bmatrix}\implies\begin{bmatrix}a\\c\end{bmatrix}=\begin{bmatrix}2\\2\end{bmatrix}\implies a=c=2.$$

Using these values,

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}1\\3\end{bmatrix}\implies\begin{bmatrix}a+b\\c+d\end{bmatrix}=\begin{bmatrix}1\\3\end{bmatrix}\implies b=-1,\quad d=1.$$

And so the desired matrix is

$$\begin{bmatrix}2 & -1\\ 2 & 1\end{bmatrix}.$$

◀

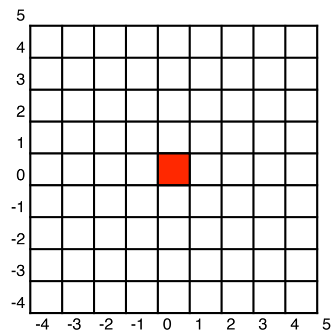

Figure 1.28: Example 51.

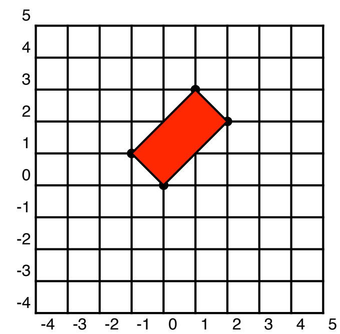

Figure 1.29: Example 51.

Homework

Problem 1.4.1 If $A=\begin{bmatrix}1 & -1\\ 2 & 3\end{bmatrix}$, $B=\begin{bmatrix}a & b\\ 1 & -2\end{bmatrix}$ and

$$(A+B)^2=A^2+2AB+B^2,$$

find a and b.

Problem 1.4.2 Consider $\triangle ABC$ with $A=\begin{pmatrix}-1\\2\end{pmatrix}$, $B=\begin{pmatrix}0\\-2\end{pmatrix}$, $C=\begin{pmatrix}2\\1\end{pmatrix}$, as in figure 1.30. Determine the the effects of the following scaling transformations on the triangle: $S_{2,1}$, $S_{1,2}$, and $S_{2,2}$.

Problem 1.4.3 Find the effects of the reflexions $R_{\frac{\pi}{2}}$, $R_{\frac{\pi}{4}}$, $R_{-\frac{\pi}{2}}$, and $R_{-\frac{\pi}{4}}$ on the triangle in figure 1.30.

Problem 1.4.4 Prove that the composition of two reflexions is commutative. Prove that the composition of a reflexion and a scaling is commutative. Prove that the composition of a reflexion and a translation is not necessarily commutative.

Problem 1.4.5 Prove that the composition of two rotations on the plane R_θ and $R_{\theta'}$ satisfies

$$R_\theta\circ R_{\theta'}=R_{\theta+\theta'}=R_{\theta'}\circ R_\theta,$$

and so the composition of two rotations on the plane is commutative.

Problem 1.4.6 Prove that the composition of a scaling and a rotation is not necessarily commutative. Prove that the composition of a rotation and a translation is not necessarily commutative. Prove that the composition of a reflexion and a rotation is not necessarily commutative.

Problem 1.4.7 Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ and $L' : \mathbb{R}^2 \to \mathbb{R}^2$ be linear transformations. Prove that their composition $L \circ L'$ is also a linear transformation.

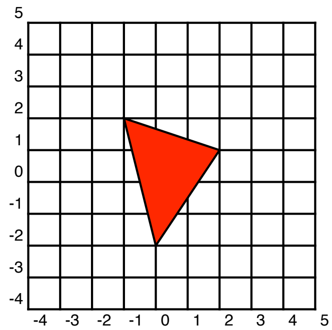

Figure 1.30: Problems 1.4.2, 1.4.8, and 1.4.3.

Problem 1.4.8 Find the effects of the reflexions R_H, R_V, and R_O on the triangle in figure 1.30.

Problem 1.4.9 Find all matrices $A \in \mathrm{M}_{2\times 2}(\mathbb{R})$ such that $A^2 = 0_2$

Problem 1.4.10 Find the image of the figure below (consisting of two circles and a triangle) under the matrix $\begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$.

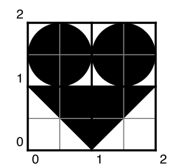

Figure 1.31: Problem 1.4.10.

1.5 Determinants in two dimensions

We now desire to define a way of determining areas of plane figures on the plane. It seems reasonable to require that this area determination agrees with common formulæ of areas of plane figures, in particular, the area of a parallelogram should be as we learn in elementary geometry and the area of a unit square should be 1.

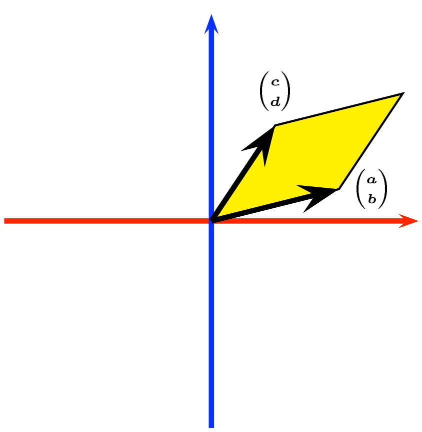

Figure 1.32: Area of a parallelogram.

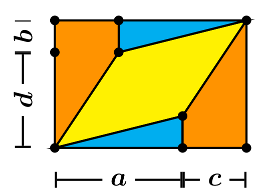

Figure 1.33: $(a+c)(b+d) - 2 \cdot \frac{ab}{2} - 2 \cdot \frac{c(2b+d)}{2} = ad - bc$.

From figures (1.32) and (1.33), the area of a parallelogram spanned by $\begin{bmatrix} a \\ b \end{bmatrix}$, and $\begin{bmatrix} c \\ d \end{bmatrix}$ is

$$D\left(\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} c \\ d \end{bmatrix}\right) = ad - bc.$$

This motivates the following definition.

52 Definition The *determinant* of the 2×2 matrix $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$ is

$$\det \begin{bmatrix} a & c \\ b & d \end{bmatrix} = ad - bc.$$

Consider now a simple quadrilateral with vertices $\mathrm{r}_1 = (x_1, y_1)$, $\mathrm{r}_2 = (x_2, y_2)$, $\mathrm{r}_3 = (x_3, y_3)$, $\mathrm{r}_4 = (x_4, y_4)$, listed in counterclockwise order, as in figure 1.34. This quadrilateral is spanned by the vectors

$$\overrightarrow{\mathrm{r}_1\mathrm{r}_2} = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix}, \qquad \overrightarrow{\mathrm{r}_1\mathrm{r}_4} = \begin{bmatrix} x_4 - x_1 \\ y_4 - y_1 \end{bmatrix},$$

and hence, its area is given by

$$A = \det \begin{bmatrix} x_2 - x_1 & x_4 - x_1 \\ y_2 - y_1 & y_4 - y_1 \end{bmatrix} = D(\overrightarrow{\mathrm{r}_2} - \overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_4} - \overrightarrow{\mathrm{r}_1}).$$

Similarly, noticing that the quadrilateral is also spanned by

$$\overrightarrow{\mathrm{r}_3\mathrm{r}_4} = \begin{bmatrix} x_4 - x_3 \\ y_4 - y_3 \end{bmatrix}, \qquad \overrightarrow{\mathrm{r}_3\mathrm{r}_2} = \begin{bmatrix} x_2 - x_3 \\ y_2 - y_3 \end{bmatrix},$$

its area is also given by

$$A = \det \begin{bmatrix} x_4 - x_3 & x_2 - x_3 \\ y_4 - y_3 & y_2 - y_3 \end{bmatrix} = D(\overrightarrow{\mathrm{r}_4} - \overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_2} - \overrightarrow{\mathrm{r}_3}).$$

Using the properties derived in Theorem **??**, we see that

$$\begin{aligned} A &= \tfrac{1}{2}\left(D(\overrightarrow{\mathrm{r}_2} - \overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_4} - \overrightarrow{\mathrm{r}_1}) + D(\overrightarrow{\mathrm{r}_4} - \overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_2} - \overrightarrow{\mathrm{r}_3})\right) \\ &= \tfrac{1}{2}\left(D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}_4}) - D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}_1}) - D(\overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_4}) + D(\overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_1})\right) \\ &\quad + \tfrac{1}{2}\left(D(\overrightarrow{\mathrm{r}_4}, \overrightarrow{\mathrm{r}_2}) - D(\overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_2}) - D(\overrightarrow{\mathrm{r}_4}, \overrightarrow{\mathrm{r}_3}) + D(\overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_3})\right) \\ &= \tfrac{1}{2}\left(D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}_4}) - D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}_1}) - D(\overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_4}))\right) + \tfrac{1}{2}\left(D(\overrightarrow{\mathrm{r}_4}, \overrightarrow{\mathrm{r}_2}) - D(\overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_2}) - D(\overrightarrow{\mathrm{r}_4}, \overrightarrow{\mathrm{r}_3})\right) \\ &= \tfrac{1}{2}\left(D(\overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_2}) + D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}_3}) + D(\overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_4}) + D(\overrightarrow{\mathrm{r}_4}, \overrightarrow{\mathrm{r}_1})\right). \end{aligned}$$

We conclude that the area of a quadrilateral with vertices (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), listed in counterclockwise order is

$$\frac{1}{2}\left(\det \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} + \det \begin{bmatrix} x_2 & x_3 \\ y_2 & y_3 \end{bmatrix} + \det \begin{bmatrix} x_3 & x_4 \\ y_3 & y_4 \end{bmatrix} + \det \begin{bmatrix} x_4 & x_1 \\ y_4 & y_1 \end{bmatrix}\right). \qquad (1.11)$$

Similarly, to find the area of a triangle of vertices $\overrightarrow{\mathrm{r}_1} = (x_1, y_1)$, $\overrightarrow{\mathrm{r}_2} = (x_2, y_2)$, $\overrightarrow{\mathrm{r}_3} = (x_3, y_3)$, listed in counterclockwise order, as in figure 1.35, reflect it about one of its sides, as in figure 1.36, creating a parallelogram. The area of the triangle is now half the area of the parallelogram, which, by virtue of 1.11, is

$$\frac{1}{4}\left(D(\overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_2}) + D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}'}) + D(\overrightarrow{\mathrm{r}'}, \overrightarrow{\mathrm{r}_3}) + D(\overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_1})\right).$$

This is equivalent to

$$\frac{1}{2}\left(D(\overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_2}) + D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}_3}) + D(\overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_1})\right) - \frac{1}{4}\left(D(\overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_2}) - D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}'}) - D(\overrightarrow{\mathrm{r}'}, \overrightarrow{\mathrm{r}_3}) + D(\overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_1}) + 2D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}_3})\right).$$

We will prove that

$$D(\overrightarrow{\mathrm{r}_1}, \overrightarrow{\mathrm{r}_2}) - D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}'}) - D(\overrightarrow{\mathrm{r}'}, \overrightarrow{\mathrm{r}_3}) + D(\overrightarrow{\mathrm{r}_3}, \overrightarrow{\mathrm{r}_1}) + 2D(\overrightarrow{\mathrm{r}_2}, \overrightarrow{\mathrm{r}_3}) = 0.$$

To do this, we appeal once again to the bi-linearity properties derived in Theorem **??**, and observe, that since we have a parallelogram, $\overrightarrow{r} - \overrightarrow{r_3} = \overrightarrow{r_2} - \overrightarrow{r_1}$, which means $\overrightarrow{r} = \overrightarrow{r_3} + \overrightarrow{r_2} - \overrightarrow{r_1}$. Thus

$$\begin{aligned}
D(\overrightarrow{r_1}, \overrightarrow{r_2}) - D(\overrightarrow{r_2}, \overrightarrow{r}) - D(\overrightarrow{r}, \overrightarrow{r_3}) + D(\overrightarrow{r_3}, \overrightarrow{r_1}) + 2D(\overrightarrow{r_2}, \overrightarrow{r_3}) &= D(\overrightarrow{r_1}, \overrightarrow{r_2}) - D(\overrightarrow{r_2}, \overrightarrow{r_3} + \overrightarrow{r_2} - \overrightarrow{r_1}) + 2D(\overrightarrow{r_2}, \overrightarrow{r_3}) \\
&\quad -D(\overrightarrow{r_3} + \overrightarrow{r_2} - \overrightarrow{r_1}, \overrightarrow{r_3}) + D(\overrightarrow{r_3}, \overrightarrow{r_1}) \\
&= D(\overrightarrow{r_1}, \overrightarrow{r_2} - \overrightarrow{r_3}) + D(\overrightarrow{r_3} + \overrightarrow{r_2} - \overrightarrow{r_1}, \overrightarrow{r_2} - \overrightarrow{r_3}) \\
&\quad +2D(\overrightarrow{r_2}, \overrightarrow{r_3}) \\
&= D(\overrightarrow{r_3} + \overrightarrow{r_2}, \overrightarrow{r_2} - \overrightarrow{r_3}) + 2D(\overrightarrow{r_2}, \overrightarrow{r_3}) \\
&= D(\overrightarrow{r_3}, \overrightarrow{r_2}) - D(\overrightarrow{r_2}, \overrightarrow{r_3}) + 2D(\overrightarrow{r_2}, \overrightarrow{r_3}) \\
&= D(\overrightarrow{r_3}, \overrightarrow{r_2}) - D(\overrightarrow{r_2}, \overrightarrow{r_3}) \\
&= 0,
\end{aligned}$$

as claimed. We have proved then that the area of a triangle, whose vertices (x_1, y_1), (x_2, y_2), (x_3, y_3) are listed in counterclockwise order, is

$$\frac{1}{2}\left(\det\begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} + \det\begin{bmatrix} x_2 & x_3 \\ y_2 & y_3 \end{bmatrix} + \det\begin{bmatrix} x_3 & x_1 \\ y_3 & y_1 \end{bmatrix}\right). \tag{1.12}$$

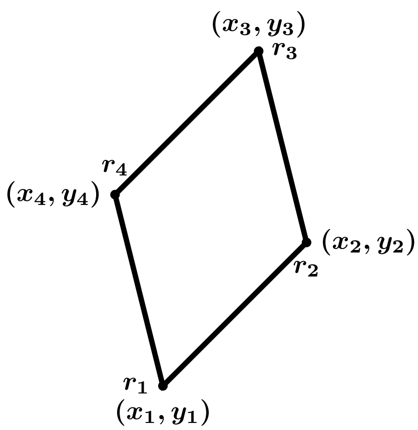

Figure 1.34: Area of a quadrilateral.

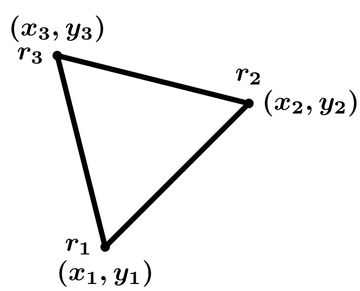

Figure 1.35: Area of a triangle.

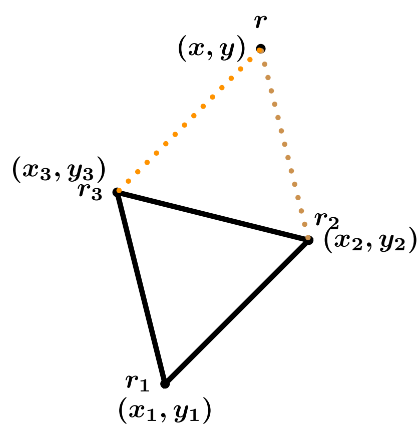

Figure 1.36: Area of a triangle.

In general, we have the following theorem.

53 Theorem (Surveyor's Theorem) Let (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n) be the vertices of a simple (non-crossing) polygon, listed in counterclockwise order. Then its area is given by

$$\frac{1}{2}\left(\det\begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} + \det\begin{bmatrix} x_2 & x_3 \\ y_2 & y_3 \end{bmatrix} + \cdots + \det\begin{bmatrix} x_{n-1} & x_n \\ y_{n-1} & y_n \end{bmatrix} + \det\begin{bmatrix} x_n & x_1 \\ y_n & y_1 \end{bmatrix}\right).$$

Proof: *The proof is by induction on n. We have already proved the cases $n = 3$ and $n = 4$ in (1.12) and (1.11), respectively. Consider now a simple polygon P with n vertices. If P is convex then we may take any vertex and draw a line to the other vertices, triangulating the polygon, creating $n - 2$ triangles. If P is not convex, then there must be a vertex that has a reflex angle. A ray produced from this vertex must hit another vertex, creating a diagonal, otherwise the polygon would have infinite area. This diagonal divides the polygon into two sub-polygons. These two sub-polygons are either both convex or at least one is not convex. In the latter case,*

we repeat the argument, finding another diagonal and creating a new sub-polygon. Eventually, since the number of vertices is infinite, we end up triangulating the polygon. Moreover, the polygon can be triangulated in such a way that all triangles inherit the positive orientation of the original polygon but each neighbouring pair of triangles have opposite orientations. Applying (1.12) we obtain that the area is

$$\sum \det \begin{bmatrix} x_i & x_j \\ y_i & y_j \end{bmatrix},$$

where the sum is over each oriented edge. Since each diagonal occurs twice, but having opposite orientations, the terms

$$\det \begin{bmatrix} x_i & x_j \\ y_i & y_j \end{bmatrix} + \det \begin{bmatrix} x_j & x_i \\ y_j & y_i \end{bmatrix} = 0,$$

disappear from the sum and we are simply left with

$$\frac{1}{2}\left(\det \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} + \det \begin{bmatrix} x_2 & x_3 \\ y_2 & y_3 \end{bmatrix} + \cdots + \det \begin{bmatrix} x_{n-1} & x_n \\ y_{n-1} & y_n \end{bmatrix} + \det \begin{bmatrix} x_n & x_1 \\ y_n & y_1 \end{bmatrix}\right).$$

❑

We may use the software Maple™ in order to speed up computations with vectors. Most of the commands we will need are in the `linalg` package. For example, let us define two vectors, $\overrightarrow{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\overrightarrow{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and a matrix $A := \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Let us compute their dot product, find a unit vector in the direction of $\overrightarrow{a}$, and the angle between the vectors. (There must be either a colon or a semicolon at the end of each statement. The result will not display if a colon is chosen.)

```
> with(linalg):
> a:=vector([2,1]);
```

$$a := [2, 1]$$

```
> b:=vector([1,2]);
```

$$b := [1, 2]$$

```
> normalize(a);
```

$$[\frac{2\sqrt{5}}{5}, \frac{\sqrt{5}}{5}]$$

```
> dotprod(a,b);
```

$$4$$

```
> angle(a,b);
```

$$\arccos\left(\frac{4}{5}\right)$$

```
> A:=matrix([[1,2],[3,4]]);
```

$$A := \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

```
> det(A);
```

$$-2$$

Homework

Problem 1.5.1 Find all vectors $\overrightarrow{a} \in \mathbb{R}^2$ such that $\overrightarrow{a} \perp \begin{bmatrix} -3 \\ 2 \end{bmatrix}$ and $||a|| = \sqrt{13}$.

Problem 1.5.2 (Pythagorean Theorem) If $\overrightarrow{a} \perp \overrightarrow{b}$, prove that

$$||\overrightarrow{a} + \overrightarrow{b}||^2 = \left|\left|\overrightarrow{a}\right|\right|^2 + \left|\left|\overrightarrow{b}\right|\right|^2.$$

Problem 1.5.3 Let a, b be arbitrary real numbers. Prove that
$$(a^2 + b^2)^2 \leq 2(a^4 + b^4).$$

Problem 1.5.4 Let $\overrightarrow{a}$, $\overrightarrow{b}$ be fixed vectors in $\mathbb{R}^2$. Prove that if
$$\forall \overrightarrow{v} \in \mathbb{R}^2, \overrightarrow{v} \bullet \overrightarrow{a} = \overrightarrow{v} \bullet \overrightarrow{b},$$
then $\overrightarrow{a} = \overrightarrow{b}$.

Problem 1.5.5 (Polarisation Identity) Let $\overrightarrow{u}$, $\overrightarrow{v}$ be vectors in $\mathbb{R}^2$. Prove that
$$\overrightarrow{u} \bullet \overrightarrow{v} = \frac{1}{4}\left(||\overrightarrow{u} + \overrightarrow{v}||^2 - ||\overrightarrow{u} - \overrightarrow{v}||^2\right).$$

Problem 1.5.6 Consider two lines on the plane L_1 and L_2 with Cartesian equations $L_1 : y = m_1x + b_1$ and $L_2 : y = m_2x + b_1$, where $m_1 \neq 0$, $m_2 \neq 0$. Using Corollary 22, prove that $L_1 \perp L_2 \iff m_1m_2 = -1$. .

Problem 1.5.7 Find the Cartesian equation of all lines L' passing through $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ and making an angle of $\frac{\pi}{6}$ radians with the Cartesian line $L : x + y = 1$.

Problem 1.5.8 Let $\overrightarrow{v}$, $\overrightarrow{w}$, be vectors on the plane, with $\overrightarrow{w} \neq \overrightarrow{0}$. Prove that the vector $\overrightarrow{a} = \overrightarrow{v} - \frac{\overrightarrow{v} \bullet \overrightarrow{w}}{||\overrightarrow{w}||^2}\overrightarrow{w}$ is perpendicular to $\overrightarrow{w}$.

1.6 Parametric Curves on the Plane

54 Definition Let $[a; b] \subseteq \mathbb{R}$. A *parametric curve* representation $\mathbf{r}$ of a curve Γ is a function $\mathbf{r} : [a; b] \to \mathbb{R}^2$, with
$$\mathbf{r}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix},$$
and such that $\mathbf{r}([a; b]) = \Gamma$. $\mathbf{r}(a)$ is the *initial point* of the curve and $\mathbf{r}(b)$ its *terminal point*. A curve is *closed* if its initial point and its final point coincide. The *trace* of the curve $\mathbf{r}$ is the set of all images of $\mathbf{r}$, that is, Γ. If there exist $t_1 \neq t_2$ such that $\mathbf{r}(t_1) = \mathbf{r}(t_2) = \mathbf{p}$, then $\mathbf{p}$ is a *multiple point* of the curve. The curve is *simple* if its has no multiple points. A closed curve whose only multiple points are its endpoints is called a *Jordan curve*.

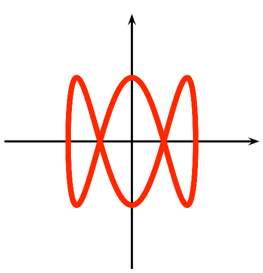

Figure 1.37: $x = \sin 2t$, $y = \cos 6t$.

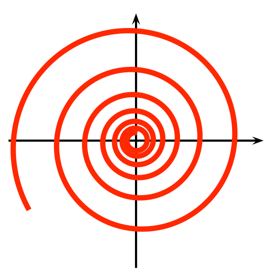

Figure 1.38: $x = 2^{t/10} \cos t$, $y = 2^{t/10} \sin t$.

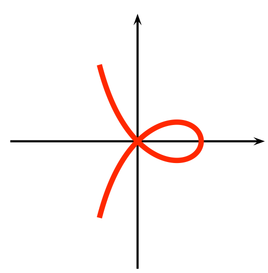

Figure 1.39: $x = \frac{1 - t^2}{1 + t^2}$, $y = \frac{t - t^3}{1 + t^2}$.

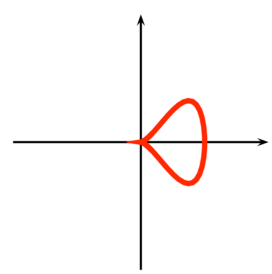

Figure 1.40: $x = (1 + \cos t)/2$, $y = (\sin t)(1 + \cos t)/2$.

Graphing parametric equations is a difficult art, and a theory akin to the one studied for Cartesian equations in a first Calculus course has been developed. Our interest is not in graphing curves, but in obtaining suitable parametrisations of simple Cartesian curves. We mention in passing however that Maple has excellent capabilities for graphing parametric equations. For example, the commands to graph the various curves in figures 1.37 through 1.40 follow.

```
> with(plots):
> plot([sin(2*t),cos(6*t),t=0..2*Pi],x=-5..5,y=-5..5);
> plot([2^(t/10)*cos(t),2^(t/10)*sin(t),t=-20..10],x=-5..5,y=-5..5);
> plot([(1-t^2)/(1+t^2),(t-t^3)/(1+t^2),t=-2..2],x=-5..5,y=-5..5);
> plot([(1+cos(t))/2,sin(t)*(1+cos(t))/2,t=0..2*Pi],x=-5..5,y=-5..5);
```

Our main focus of attention will be the following. Given a Cartesian curve with equation $f(x, y) = 0$, we wish to find suitable parametrisations for them. That is, we want to find functions $x : t \mapsto a(t)$, $y : t \mapsto b(t)$ and an interval I such that the graphs of $f(x, y) = 0$ and $f(a(t), b(t)) = 0$, $t \in I$ coincide. These parametrisations may differ in features, according to the choice of functions and the choice of intervals.

55 Example Consider the parabola with Cartesian equation $y = x^2$. We will give various parametrisations for portions of this curve.

1. If $x = t$ and $y = t^2$, then clearly $y = t^2 = x^2$. This works for every $t \in \mathbb{R}$, and hence the parametrisation
$$x = t, \quad y = t^2, \quad t \in \mathbb{R}$$
works for the whole curve. Notice that as t increases, the curve is traversed from left to right.

2. If $x = \sqrt{t}$ and $y = t$, then again $y = t = (\sqrt{t})^2 = x^2$. This works only for $t \geq 0$, and hence the parametrisation
$$x = \sqrt{t}, \quad y = t, \quad t \in [0; +\infty[$$
gives the half of the curve for which $x \geq 0$. As t increases, the curve is traversed from left to right.

3. Similarly, if $x = -\sqrt{t}$ and $y = t$, then again $y = t = (-\sqrt{t})^2 = x^2$. This works only for $t \geq 0$, and hence the parametrisation
$$x = -\sqrt{t}, \quad y = t, \quad t \in [0; +\infty[$$
gives the half of the curve for which $x \leq 0$. As t increases, x decreases, and so the curve is traversed from right to left.

4. If $x = \cos t$ and $y = \cos^2 t$, then again $y = \cos^2 t = (\cos t)^2 = x^2$. Both x and y are periodic with period 2π, and so this parametrisation only agrees with the curve $y = x^2$ when $-1 \leq x \leq 1$. For $t \in [0; \pi]$, the cosine decreases from 1 to -1 and so the curve is traversed from right to left in this interval.

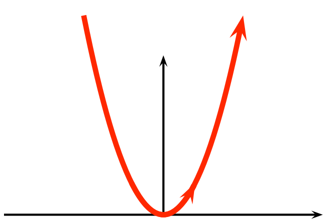

Figure 1.41: $x = t, y = t^2$, $t \in \mathbb{R}$.

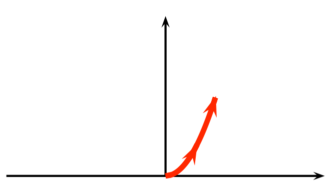

Figure 1.42: $x = \sqrt{t}, y = t$, $t \in [0; +\infty[$.

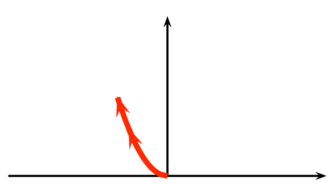

Figure 1.43: $x = -\sqrt{t}, y = t$, $t \in [0; +\infty[$.

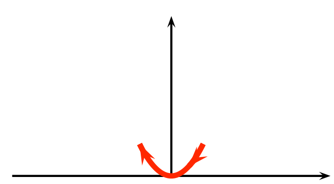

Figure 1.44: $x = \cos t, y = \cos^2 t$, $t \in [0; \pi]$.

The identities
$$\cos^2\theta + \sin^2\theta = 1, \qquad \tan^2\theta - \sec^2\theta = 1, \qquad \cosh^2\theta - \sinh^2\theta = 1,$$
are often useful when parametrising quadratic curves.

56 Example Give two distinct parametrisations of the ellipse $\dfrac{(x-1)^2}{4} + \dfrac{(y+2)^2}{9} = 1$.

1. The first parametrisation must satisfy that as t traverses the values in the interval $[0; 2\pi]$, one starts at the point $(3, -2)$, traverses the ellipse once counterclockwise, finishing at $(3, -2)$.

2. The second parametrisation must satisfy that as t traverses the interval $[0; 1]$, one starts at the point $(3, -2)$, traverses the ellipse twice clockwise, and returns to $(3, -2)$.

Solution: ▶ *What formula do we know where a sum of two squares equals 1? We use a trigonometric substitution, a sort of "polar coordinates." Observe that for $t \in [0; 2\pi]$, the point $(\cos t, \sin t)$ traverses the unit circle once, starting at $(1, 0)$ and ending there. Put*

$$\frac{x-1}{2} = \cos t \implies x = 1 + 2\cos t,$$

and

$$\frac{y+2}{3} = \sin t \implies y = -2 + 3\sin t.$$

Then

$$x = 1 + 2\cos t, \quad y = -2 + 3\sin t, \quad t \in [0; 2\pi]$$

is the desired first parametrisation.

For the second parametrisation, notice that as t traverses the interval $[0; 1]$, $(\sin 4\pi t, \cos 4\pi t)$ traverses the unit interval twice, clockwise, but begins and ends at the point $(0, 1)$. To begin at the point $(1, 0)$ we must make a shift: $\left(\sin\left(4\pi t + \frac{\pi}{2}\right), \cos\left(4\pi t + \frac{\pi}{2}\right)\right)$ will start at $(1, 0)$ and travel clockwise twice, as t traverses $[0; 1]$. Hence we may take

$$x = 1 + 2\sin\left(4\pi t + \frac{\pi}{2}\right), \quad y = -2 + 3\cos\left(4\pi t + \frac{\pi}{2}\right), \quad t \in [0; 1]$$

as our parametrisation. ◀

Some classic curves can be described by mechanical means, as the curves drawn by a spirograph. We will consider one such curve.

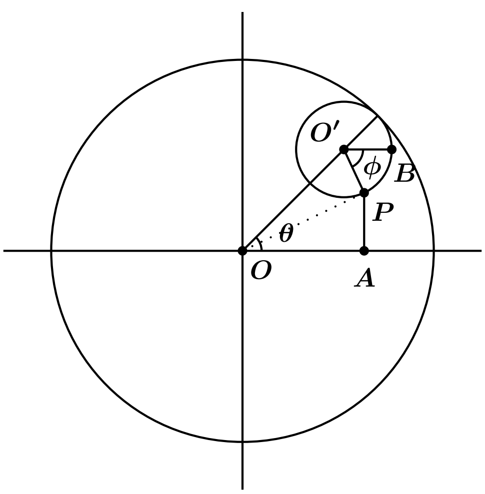

Figure 1.45: Construction of the hypocycloid.

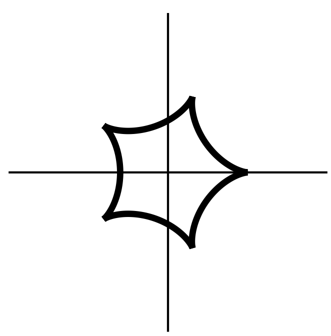

Figure 1.46: Hypocycloid with $R = 5$, $\rho = 1$.

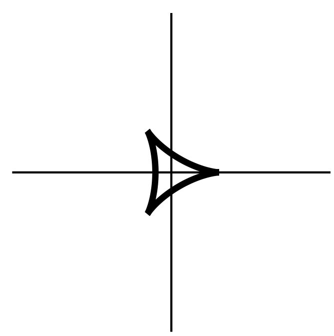

Figure 1.47: Hypocycloid with $R = 3$, $\rho = 2$.

57 Example A *hypocycloid* is a curve traced out by a fixed point P on a circle $\mathscr{C}$ of radius ρ as $\mathscr{C}$ rolls on the inside of a circle with centre at O and radius R. If the initial position of P is $\begin{pmatrix} R \\ 0 \end{pmatrix}$, and θ is the angle, measured counterclockwise, that a ray starting at O and passing through the centre of $\mathscr{C}$ makes with the x-axis, shew that a parametrisation of the hypocycloid is

$$x = (R - \rho)\cos\theta + \rho\cos\left(\frac{(R-\rho)\theta}{\rho}\right),$$

$$y = (R-\rho)\sin\theta - \rho\sin\left(\frac{(R-\rho)\theta}{\rho}\right).$$

Solution: ► *Suppose that starting from $\theta = 0$, the centre O' of the small circle moves counterclockwise inside the larger circle by an angle θ, and the point $P = (x, y)$ moves clockwise an angle ϕ. The arc length travelled by the centre of the small circle is $(R-\rho)\theta$ radians. At the same time the point P has rotated $\rho\phi$ radians, and so $(R-\rho)\theta = \rho\phi$. See figure 1.45, where $O'B$ is parallel to the x-axis.*

Let A be the projection of P on the x-axis. Then $\angle OAP = \angle OPO' = \frac{\pi}{2}$, $\angle OO'P = \pi - \phi - \theta$, $\angle POA = \frac{\pi}{2} - \phi$, and $OP = (R-\rho)\sin(\pi - \phi - \theta)$. Hence

$$x = (OP)\cos\angle POA = (R-\rho)\sin(\pi-\phi-\theta)\cos(\frac{\pi}{2}-\phi),$$

$$y = (R-\rho)\sin(\pi-\phi-\theta)\sin(\frac{\pi}{2}-\phi).$$

Now

$$\begin{aligned} x &= (R-\rho)\sin(\pi-\phi-\theta)\cos(\frac{\pi}{2}-\phi) \\ &= (R-\rho)\sin(\phi+\theta)\sin\phi \\ &= \frac{(R-\rho)}{2}(\cos\theta - \cos(2\phi+\theta)) \\ &= (R-\rho)\cos\theta - \frac{(R-\rho)}{2}(\cos\theta + \cos(2\phi+\theta)) \\ &= (R-\rho)\cos\theta - (R-\rho)(\cos(\theta+\phi)\cos\phi). \end{aligned}$$

Also, $\cos(\theta+\phi) = -\cos(\pi-\theta-\phi) = -\frac{\rho}{OO'} = -\frac{\rho}{R-\rho}$ and $\cos\phi = \cos\left(\frac{(R-\rho)\theta}{\rho}\right)$ and so

$$x = (R-\rho)\cos\theta - (R-\rho)(\cos(\theta+\phi)\cos\phi = (R-\rho)\cos\theta + \rho\cos\left(\frac{(R-\rho)\theta}{\rho}\right),$$

as required. The identity for y is proved similarly. A particular example appears in figure 1.47. ◄

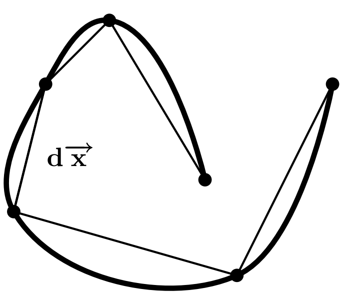

Figure 1.48: Length of a curve.

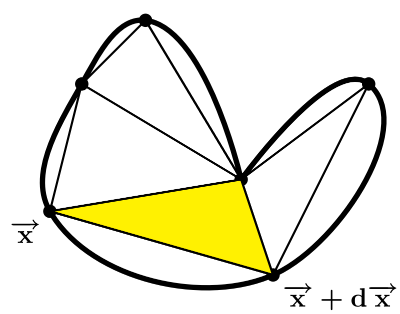

Figure 1.49: Area enclosed by a simple closed curve

Given a curve Γ how can we find its length? The idea, as seen in figure 1.48 is to consider the projections $\mathrm{d}x, \mathrm{d}y$ at each point. The length of the vector

$$\mathrm{dr} = \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix}$$

is

$$||\mathbf{dr}|| = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2}.$$

Hence the length of $\mathbf{\Gamma}$ is given by

$$\int_{\Gamma} ||\mathbf{dr}|| = \int_{\Gamma} \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2}. \tag{1.13}$$

Similarly, suppose that $\mathbf{\Gamma}$ is a simple closed curve in $\mathbb{R}^2$. How do we find the (oriented) area of the region it encloses? The idea, borrowed from finding areas of polygons, is to split the region into triangles, each of area

$$\frac{1}{2}\det\begin{bmatrix} x & x+\mathrm{d}x \\ y & y+\mathrm{d}y \end{bmatrix} = \frac{1}{2}\det\begin{bmatrix} x & \mathrm{d}x \\ y & \mathrm{d}y \end{bmatrix} = \frac{1}{2}(x\mathrm{d}y - y\mathrm{d}x),$$

and to sum over the closed curve, obtaining a total oriented area of

$$\frac{1}{2}\oint_{\Gamma}\det\begin{bmatrix} x & \mathrm{d}x \\ y & \mathrm{d}y \end{bmatrix} = \frac{1}{2}\oint_{\Gamma}(x\mathrm{d}y - y\mathrm{d}x). \tag{1.14}$$

Here $\oint_{\Gamma}$ denotes integration around the closed curve.

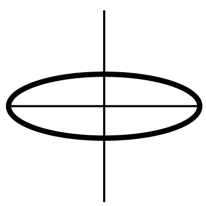

Figure 1.50: Example 58.

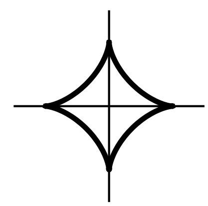

Figure 1.51: Example 59.

58 Example Let $(A, B) \in \mathbb{R}^2, A > 0, B > 0$. Find a parametrisation of the ellipse

$$\Gamma : \left\{(x, y) \in \mathbb{R}^2 : \frac{x^2}{A^2} + \frac{y^2}{B^2} = 1\right\}.$$

Furthermore, find an integral expression for the perimeter of this ellipse and find the area it encloses.

Solution: ► *Consider the parametrisation* $\Gamma : [0; 2\pi] \to \mathbb{R}^2$, *with*

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A\cos t \\ B\sin t \end{bmatrix}.$$

This is a parametrisation of the ellipse, for

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = \frac{A^2\cos^2 t}{A^2} + \frac{B^2\sin^2 t}{B^2} = \cos^2 t + \sin^2 t = 1.$$

Notice that this parametrisation goes around once the ellipse counterclockwise. The perimeter of the ellipse is given by

$$\int_{\Gamma} ||\mathrm{d}\,\overrightarrow{\mathrm{r}}|| = \int_0^{2\pi} \sqrt{A^2\sin^2 t + B^2\cos^2 t}\,\mathrm{d}t.$$

The above integral is an elliptic integral, *and we do not have a closed form for it (in terms of the elementary functions studied in Calculus I). We will have better luck with the area of the ellipse, which is given by*

$$\begin{aligned}\frac{1}{2}\oint_\Gamma (x\mathrm{d}y - y\mathrm{d}x) &= \frac{1}{2}\oint (A\cos t\ \mathrm{d}(B\sin t) - B\sin t\ \mathrm{d}(A\cos t))\\ &= \frac{1}{2}\int_0^{2\pi} (AB\cos^2 t + AB\sin^2 t))\ \mathrm{d}t\\ &= \frac{1}{2}\int_0^{2\pi} AB\ \mathrm{d}t\\ &= \pi AB.\end{aligned}.$$

◄

59 Example Find a parametric representation for the astroid

$$\Gamma : \left\{(x,y)\in\mathbb{R}^2 : x^{2/3}+y^{2/3}=1\right\},$$

in figure 1.51. Find the perimeter of the astroid and the area it encloses.

Solution: ► *Take*

$$\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}\cos^3 t\\ \sin^3 t\end{bmatrix}$$

with $t\in[0;2\pi]$. *Then*

$$x^{2/3}+y^{2/3} = \cos^2 t + \sin^2 t = 1.$$

The perimeter of the astroid is

$$\begin{aligned}\int_\Gamma ||\mathrm{d}\mathbf{r}|| &= \int_0^{2\pi}\sqrt{9\cos^4 t\sin^2 t + 9\sin^4 t\cos^2 t}\ \mathrm{d}t\\ &= \int_0^{2\pi} 3|\sin t\cos t|\ \mathrm{d}t\\ &= \frac{3}{2}\int_0^{2\pi}|\sin 2t|\ \mathrm{d}t\\ &= 6\int_0^{\pi/2}\sin 2t\ \mathrm{d}t\\ &= 6.\end{aligned}$$

The area of the astroid is given by

$$\begin{aligned}\frac{1}{2}\oint_\Gamma (x\mathrm{d}y - y\mathrm{d}x) &= \frac{1}{2}\oint (\cos^3 t\ \mathrm{d}(\sin^3 t) - \sin^3 t\ \mathrm{d}(\cos^3 t))\\ &= \frac{1}{2}\int_0^{2\pi}(3\cos^4 t\sin^2 t + 3\sin^4 t\cos^2 t))\ \mathrm{d}t\\ &= \frac{3}{2}\int_0^{2\pi}(\sin t\cos t)^2\ \mathrm{d}t\\ &= \frac{3}{8}\int_0^{2\pi}(\sin 2t)^2\ \mathrm{d}t\\ &= \frac{3}{16}\int_0^{2\pi}(1-\cos 4t)\ \mathrm{d}t\\ &= \frac{3\pi}{8}.\end{aligned}$$

We can use Maple™ (at least version 10) to calculate the above integrals. For example, if $(x,y) = (\cos^3 t, \sin^3 t)$, *to compute the arc length we use the path integral command and to compute the area, we use the line integral command with the vector field* $\begin{bmatrix}-y/2\\ x/2\end{bmatrix}$.

```
> with(Student[VectorCalculus]):
> PathInt( 1, [x,y]=Path( <(cos(t))^3,(sin(t))^3>,0..2*Pi));
> LineInt( VectorField(<-y/2,x/2>), Path( <(cos(t))^3,(sin(t))^3>,0..2*Pi));
```

◀

We include here for convenience, some Maple commands to compute various arc lengths and areas.

60 Example To obtain the arc length of the path in figure 1.52, we type

```
> with(Student[VectorCalculus]):
> PathInt(1, [x,y] = LineSegments( <0,0>, <1,1>, <1,2> ,<2,1>,<3,3>,<4,1>);
```

To obtain the arc length of the path in figure 1.53, we type

```
> with(Student[VectorCalculus]):
> PathInt(1, [x,y] =  Arc( Circle( <0,0>, 3), Pi/6, Pi/5 ) ;
```

To obtain the area inside the curve in 1.54

```
> with(Student[VectorCalculus]):
> LineInt( VectorField(<-y/2,x/2>),
> Path( <(1+cos(t))*(cos(t))+1,(1+cos(t))*(sin(t))+2>,0..2*Pi));
```

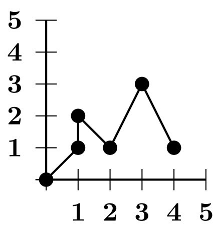

Figure 1.52: Line Path.

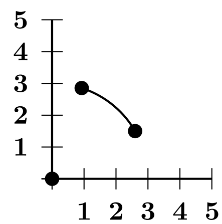

Figure 1.53: Arc of circle of radius 3, angle $\frac{\pi}{6} \leq \theta \leq \frac{\pi}{5}$.

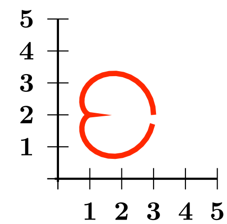

Figure 1.54: $x = 1 + (1 + \cos t)(\cos t), y = 2 + (1 + \cos t)(\sin t)$.

Homework

Problem 1.6.1 A curve is represented parametrically by $x(t) = t^3 - 2t$, $y(t) = t^3 + 2t$. Find its Cartesian equation.

Problem 1.6.2 Give an implicit Cartesian equation for the parametric representation $x = \dfrac{t^2}{1+t^5}$, $y = \dfrac{t^3}{1+t^5}$.

Problem 1.6.3 Let a, b, c, d be strictly positive real constants. In each case give an implicit Cartesian equation for the parametric representation and describe the trace of the parametric curve.

1. $x = at + b$, $y = ct + d$
2. $x = \cos t$, $y = 0$
3. $x = a \cosh t$, $y = b \sinh t$
4. $x = a \sec t$, $y = b \tan t$, $t \in]-\frac{\pi}{2}; \frac{\pi}{2}[$

Problem 1.6.4 Parametrise the curve $y = \log \cos x$ for $0 \leq x \leq \frac{\pi}{3}$. Then find its arc length.

Problem 1.6.5 Describe the trace of the parametric curve

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \sin t \\ 2\sin t + 1 \end{bmatrix}, \quad t \in [0; 4\pi].$$

Problem 1.6.6 Consider the plane curve defined implicitly by $\sqrt{x} + \sqrt{y} = 1$. Give a suitable parametrisation of this curve, and find its length. The graph of the curve appears in figure 1.55.

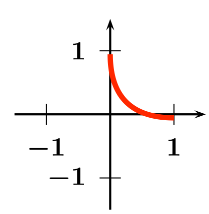

Figure 1.55: Problem 1.6.6.

Problem 1.6.7 Consider the graph given parametrically by $x(t) = t^3 + 1$, $y(t) = 1 - t^2$. Find the area under the graph, over the x axis, and between the lines $x = 1$ and $x = 2$.

Problem 1.6.8 Find the arc length of the curve given parametrically by $x(t) = 3t^2$, $y(t) = 2t^3$ for $0 \leq t \leq 1$.

Problem 1.6.9 Let $\mathscr{C}$ be the curve in $\mathbb{R}^2$ defined by

$$x(t) = \frac{t^2}{2}, \quad y(t) = \frac{(2t+1)^{3/2}}{3}, \quad t \in [-\frac{1}{2}; +\frac{1}{2}].$$

Find the length of this curve.

Problem 1.6.10 Find the area enclosed by the curve $x(t) = \sin^3 t, y(t) = (\cos t)(1 + \sin^2 t)$. The curve appears in figure 1.56.

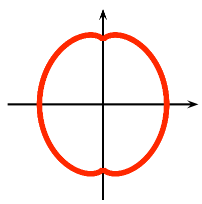

Figure 1.56: Problem 1.6.10.

Problem 1.6.11 Let $\mathscr{C}$ be the curve in $\mathbb{R}^2$ defined by

$$x(t) = \frac{3t}{1+t^3}, \quad y(t) = \frac{3t^2}{1+t^3}, \quad t \in \mathbb{R} \setminus \{-1\},$$

which you may see in figure 1.57. Find the area enclosed by the loop of this curve.

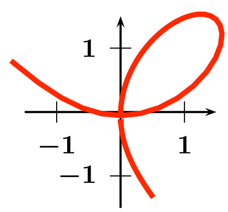

Figure 1.57: Problem 1.6.11.

Problem 1.6.12 Let P be a point at a distance d from the centre of a circle of radius ρ. The curve traced out by P as the circle rolls along a straight line, without slipping, is called a *cycloid*. Find a parametrisation of the cycloid.

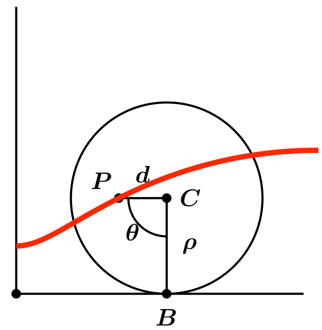

Figure 1.58: Cycloid

Problem 1.6.13 Find the arc length of the arc of the cycloid $x = \rho(t - \cos t)$, $y = \rho(1 - \cos t)$, $t \in [0; 2\pi]$.

Problem 1.6.14 Find the length of the parametric curve given by

$$x = e^t \cos t, \qquad y = e^t \sin t, \qquad t \in [0; \pi].$$

Problem 1.6.15 A shell strikes an airplane flying at height h above the ground. It is known that the shell was fired from a gun on the ground with a muzzle velocity of magnitude V, but the position of the gun and its angle of elevation are both unknown. Deduce that the gun is situated within a circle whose centre lies directly below the airplane and whose radius is

$$\frac{V\sqrt{V^2 - 2gh}}{g}.$$

Problem 1.6.16 The parabola $y^2 = -4px$ rolls without slipping around the parabola $y^2 = 4px$. Find the equation of the locus of the vertex of the rolling parabola.

1.7 Vectors in Space

61 Definition The 3-dimensional Cartesian Space is defined and denoted by

$$\mathbb{R}^3 = \{\mathbf{r} = (x, y, z) : x \in \mathbb{R}, \ y \in \mathbb{R}, \ z \in \mathbb{R}\}.$$

In figure 1.59 we have pictured the point $(2, 1, 3)$.

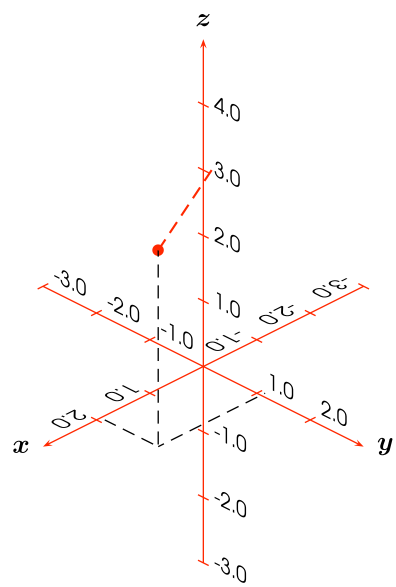

Figure 1.59: A point in space.

Having oriented the z axis upwards, we have a choice for the orientation of the the x and y-axis. We adopt a convention known as a *right-handed coordinate system*, as in figure 1.60. Let us explain. In analogy to $\mathbb{R}^2$ we put

$$\overrightarrow{\mathrm{i}} = \begin{bmatrix}1\\0\\0\end{bmatrix}, \qquad \overrightarrow{\mathrm{j}} = \begin{bmatrix}0\\1\\0\end{bmatrix}, \qquad \overrightarrow{\mathrm{k}} = \begin{bmatrix}0\\0\\1\end{bmatrix},$$

and observe that

$$\mathrm{r} = (x, y, z) = x\,\overrightarrow{\mathrm{i}} + y\,\overrightarrow{\mathrm{j}} + z\,\overrightarrow{\mathrm{k}}.$$

Most of what we did in $\mathbb{R}^2$ transfers to $\mathbb{R}^3$ without major complications.

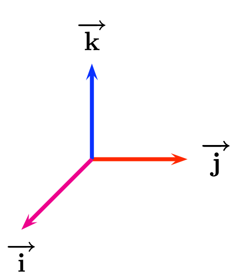

Figure 1.60: Right-handed system.

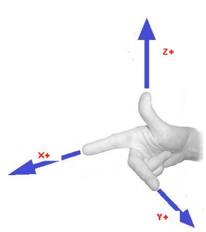

Figure 1.61: Right Hand.

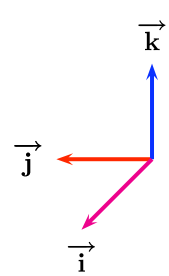

Figure 1.62: Left-handed system.

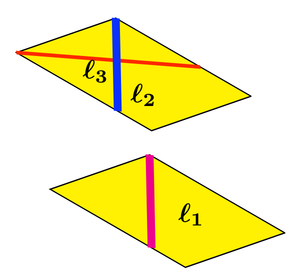

Figure 1.63: $\ell_1 \parallel \ell_2$. ℓ_1 and ℓ_3 are skew.

62 Definition The dot product of two vectors $\overrightarrow{a}$ and $\overrightarrow{b}$ in $\mathbb{R}^3$ is

$$\overrightarrow{a}\bullet\overrightarrow{b} = a_1b_1 + a_2b_2 + a_3b_3.$$

The norm of a vector $\overrightarrow{a}$ in $\mathbb{R}^3$ is

$$\left|\left|\overrightarrow{a}\right|\right| = \sqrt{\overrightarrow{a}\bullet\overrightarrow{a}} = \sqrt{(a_1)^2 + (a_2)^2 + (a_3)^2}.$$

Just as in $\mathbb{R}^2$, the dot product satisfies $\overrightarrow{a}\bullet\overrightarrow{b} = \left|\left|\overrightarrow{a}\right|\right|\left|\left|\overrightarrow{b}\right|\right|\cos\theta$, where $\theta \in [0;\pi]$ is the convex angle between the two vectors.

The Cauchy-Schwarz-Bunyakovsky Inequality takes the form

$$|\overrightarrow{a}\bullet\overrightarrow{b}| \leq \left|\left|\overrightarrow{a}\right|\right|\left|\left|\overrightarrow{b}\right|\right| \implies |a_1b_1 + a_2b_2 + a_3b_3| \leq (a_1^2 + a_2^2 + a_3^2)^{1/2}(b_1^2 + b_2^2 + b_3^2)^{1/2},$$

equality holding if an only if the vectors are parallel.

63 Example Let x, y, z be positive real numbers such that $x^2 + 4y^2 + 9z^2 = 27$. Maximise $x + y + z$.

Solution: ► *Since x, y, z are positive, $|x + y + z| = x + y + z$. By Cauchy's Inequality,*

$$|x+y+z| = \left|x + 2y\left(\frac{1}{2}\right) + 3z\left(\frac{1}{3}\right)\right| \leq (x^2+4y^2+9z^2)^{1/2}\left(1 + \frac{1}{4} + \frac{1}{9}\right)^{1/2} = \sqrt{27}\left(\frac{7}{6}\right) = \frac{7\sqrt{3}}{2}.$$

Equality occurs if and only if

$$\begin{bmatrix} x \\ 2y \\ 3z \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \end{bmatrix} \implies x = \lambda, y = \frac{\lambda}{4}, z = \frac{\lambda}{9} \implies \lambda^2 + \frac{\lambda^2}{4} + \frac{\lambda^2}{9} = 27 \implies \lambda = \pm\frac{18\sqrt{3}}{7}.$$

Therefore for a maximum we take

$$x = \frac{18\sqrt{3}}{7}, \quad y = \frac{9\sqrt{3}}{14}, \quad z = \frac{2\sqrt{3}}{7}.$$

◄

64 Definition Let a be a point in $\mathbb{R}^3$ and let $\overrightarrow{v} \neq \overrightarrow{0}$ be a vector in $\mathbb{R}^3$. The *parametric line* passing through a in the direction of $\overrightarrow{v}$ is the set

$$\left\{r \in \mathbb{R}^3 : r = a + t\overrightarrow{v}\right\}.$$

65 Example Find the parametric equation of the line passing through $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}$.

Solution: ► *The line follows the direction*

$$\begin{bmatrix} 1-(-2) \\ 2-(-1) \\ 3-0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

The desired equation is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t\begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

◄

☞ *Given two lines in space, one of the following three situations might arise: (i) the lines intersect at a point, (ii) the lines are parallel, (iii) the lines are skew (non-parallel, one over the other, without intersecting, lying on different planes). See figure 1.63.*

Consider now two non-zero vectors $\overrightarrow{\mathrm{a}}$ and $\overrightarrow{\mathrm{b}}$ in $\mathbb{R}^3$. If $\overrightarrow{\mathrm{a}} \parallel \overrightarrow{\mathrm{b}}$, then the set

$$\{s\overrightarrow{\mathrm{a}} + t\overrightarrow{\mathrm{b}} : s \in \mathbb{R}, t \in \mathbb{R}\} = \{\lambda\overrightarrow{\mathrm{a}} : \lambda \in \mathbb{R}\},$$

which is a line through the origin. Suppose now that $\overrightarrow{\mathrm{a}}$ and $\overrightarrow{\mathrm{b}}$ are not parallel. We saw in the preceding chapter that if the vectors were on the plane, they would span the whole plane $\mathbb{R}^2$. In the case at hand the vectors are in space, they still span a plane, passing through the origin. Thus

$$\{s\overrightarrow{\mathrm{a}} + t\overrightarrow{\mathrm{b}} : s \in \mathbb{R}, t \in \mathbb{R}, \overrightarrow{\mathrm{a}} \not\parallel \overrightarrow{\mathrm{b}}\}$$

is a plane passing through the origin. We will say, abusing language, that two vectors are *coplanar* if there exists bi-point representatives of the vector that lie on the same plane. We will say, again abusing language, that a vector is *parallel to a specific plane* or that it *lies on a specific plane* if there exists a bi-point representative of the vector that lies on the particular plane. All the above gives the following result.

66 Theorem Let $\overrightarrow{\mathrm{v}}, \overrightarrow{\mathrm{w}}$ in $\mathbb{R}^3$ be non-parallel vectors. Then every vector $\overrightarrow{\mathrm{u}}$ of the form

$$\overrightarrow{\mathrm{u}} = a\overrightarrow{\mathrm{v}} + b\overrightarrow{\mathrm{w}},$$

a, b arbitrary scalars, is coplanar with both $\overrightarrow{\mathrm{v}}$ and $\overrightarrow{\mathrm{w}}$. Conversely, any vector $\overrightarrow{\mathrm{t}}$ coplanar with both $\overrightarrow{\mathrm{v}}$ and $\overrightarrow{\mathrm{w}}$ can be uniquely expressed in the form

$$\overrightarrow{\mathrm{t}} = p\overrightarrow{\mathrm{v}} + q\overrightarrow{\mathrm{w}}.$$

See figure 1.64.

From the above theorem, if a vector $\overrightarrow{\mathrm{a}}$ is not a linear combination of two other vectors $\overrightarrow{\mathrm{b}}, \overrightarrow{\mathrm{c}}$, then linear combinations of these three vectors may lie outside the plane containing $\overrightarrow{\mathrm{b}}, \overrightarrow{\mathrm{c}}$. This prompts the following theorem.

67 Theorem Three vectors $\overrightarrow{\mathrm{a}}, \overrightarrow{\mathrm{b}}, \overrightarrow{\mathrm{c}}$ in $\mathbb{R}^3$ are said to be *linearly independent* if

$$\alpha\overrightarrow{\mathrm{a}} + \beta\overrightarrow{\mathrm{b}} + \gamma\overrightarrow{\mathrm{c}} = \overrightarrow{0} \implies \alpha = \beta = \gamma = 0.$$

Any vector in $\mathbb{R}^3$ can be written as a linear combination of three linearly independent vectors in $\mathbb{R}^3$.

A plane is determined by three non-collinear points. Suppose that $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are non-collinear points on the same plane and that $\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is another arbitrary point on this plane. Since $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are non-collinear, $\overrightarrow{\mathbf{ab}}$ and $\overrightarrow{\mathbf{ac}}$, which are coplanar, are non-parallel. Since $\overrightarrow{\mathbf{ax}}$ also lies on the plane, we have by Lemma 66, that there exist real numbers p, q with

$$\overrightarrow{\mathbf{ar}} = p\overrightarrow{\mathbf{ab}} + q\overrightarrow{\mathbf{ac}}.$$

By Chasles' Rule,

$$\overrightarrow{\mathrm{r}} = \overrightarrow{\mathrm{a}} + p(\overrightarrow{\mathrm{b}} - \overrightarrow{\mathrm{a}}) + q(\overrightarrow{\mathrm{c}} - \overrightarrow{\mathrm{a}}),$$

is the equation of a plane containing the three non-collinear points a, b, and c, where $\overrightarrow{a}$, $\overrightarrow{b}$, and $\overrightarrow{c}$ are the position vectors of these points. Thus we have the following theorem.

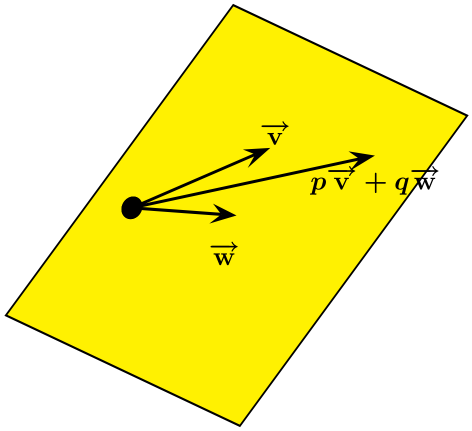

Figure 1.64: Theorem 66.

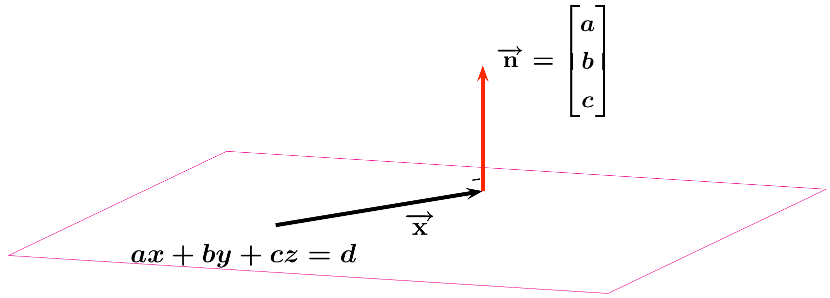

Figure 1.65: Theorem 69.

68 Theorem Let $\overrightarrow{u}$ and $\overrightarrow{v}$ be linearly independent vectors. The *parametric equation* of a plane containing the point a, and parallel to the vectors $\overrightarrow{u}$ and $\overrightarrow{v}$ is given by

$$\overrightarrow{r} - \overrightarrow{a} = p\overrightarrow{u} + q\overrightarrow{v}.$$

Componentwise this takes the form

$$\begin{aligned} x - a_1 &= pu_1 + qv_1, \\ y - a_2 &= pu_2 + qv_2, \\ z - a_3 &= pu_3 + qv_3. \end{aligned}$$

Multiplying the first equation by $u_2v_3 - u_3v_2$, the second by $u_3v_1 - u_1v_3$, and the third by $u_1v_2 - u_2v_1$, we obtain,

$$\begin{aligned} (u_2v_3 - u_3v_2)(x - a_1) &= (u_2v_3 - u_3v_2)(pu_1 + qv_1), \\ (u_3v_1 - u_1v_3)(y - a_2) &= (u_3v_1 - u_1v_3)(pu_2 + qv_2), \\ (u_1v_2 - u_2v_1)(z - a_3) &= (u_1v_2 - u_2v_1)(pu_3 + qv_3). \end{aligned}$$

Adding gives,

$$(u_2v_3 - u_3v_2)(x - a_1) + (u_3v_1 - u_1v_3)(y - a_2) + (u_1v_2 - u_2v_1)(z - a_3) = 0.$$

Put

$$a = u_2v_3 - u_3v_2, \qquad b = u_3v_1 - u_1v_3, \qquad c = u_1v_2 - u_2v_1,$$

and

$$d = a_1(u_2v_3 - u_3v_2) + a_2(u_3v_1 - u_1v_3) + a_3(u_1v_2 - u_2v_1).$$

Since $\overrightarrow{v}$ is linearly independent from $\overrightarrow{u}$, not all of a, b, c are zero. This gives the following theorem.

69 Theorem The equation of the plane in space can be written in the form

$$ax + by + cz = d,$$

which is the *Cartesian equation* of the plane. Here $a^2 + b^2 + c^2 \neq 0$, that is, at least one of the coefficients is non-zero. Moreover, the vector $\overrightarrow{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is normal to the plane with Cartesian equation $ax + by + cz = d$.

Proof: *We have already proved the first statement. For the second statement, observe that if $\overrightarrow{u}$ and $\overrightarrow{v}$ are non-parallel vectors and $\overrightarrow{r} - \overrightarrow{a} = p\overrightarrow{u} + q\overrightarrow{v}$ is the equation of the plane containing the point* a *and parallel to the vectors $\overrightarrow{u}$ and $\overrightarrow{v}$, then if $\overrightarrow{n}$ is simultaneously perpendicular to $\overrightarrow{u}$ and $\overrightarrow{v}$ then $(\overrightarrow{r} - \overrightarrow{a})\bullet\overrightarrow{n} = 0$ for $\overrightarrow{u}\bullet\overrightarrow{n} = 0 = \overrightarrow{v}\bullet\overrightarrow{n}$. Now, since at least one of a, b, c is non-zero, we may assume $a \neq 0$. The argument is similar if one of the other letters is non-zero and $a = 0$. In this case we can see that*

$$x = \frac{d}{a} - \frac{b}{a}y - \frac{c}{a}z.$$

Put $y = s$ and $z = t$. Then

$$\begin{pmatrix} x - \frac{d}{a} \\ y \\ z \end{pmatrix} = s\begin{bmatrix} -\frac{b}{a} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -\frac{c}{a} \\ 0 \\ 1 \end{bmatrix}$$

is a parametric equation for the plane. We have

$$a\left(-\frac{b}{a}\right) + b\,(1) + c\,(0) = 0, \qquad a\left(-\frac{c}{a}\right) + b\,(0) + c\,(1) = 0,$$

and so $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is simultaneously perpendicular to $\begin{bmatrix} -\frac{b}{a} \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -\frac{c}{a} \\ 0 \\ 1 \end{bmatrix}$, proving the second statement.❑

70 Example The equation of the plane passing through the point $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ and normal to the vector $\begin{bmatrix} -3 \\ 2 \\ 4 \end{bmatrix}$ is

$$-3(x-1) + 2(y+1) + 4(z-2) = 0 \implies -3x + 2y + 4z = 3.$$

71 Example Find both the parametric equation and the Cartesian equation of the plane parallel to the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and passing through the point $\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$.

Solution: ► *The desired parametric equation is*

$$\begin{pmatrix} x \\ y+1 \\ z-2 \end{pmatrix} = s\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

This gives

$$s = z - 2, \qquad t = y + 1 - s = y + 1 - z + 2 = y - z + 3$$

and

$$x = s + t = z - 2 + y - z + 3 = y + 1.$$

Hence the Cartesian equation is $x - y = 1$. ◄

72 Definition If $\overrightarrow{\mathbf{n}}$ is perpendicular to plane Π_1 and $\overrightarrow{\mathbf{n}}'$ is perpendicular to plane Π_2, the *angle between the two planes* is the angle between the two vectors $\overrightarrow{\mathbf{n}}$ and $\overrightarrow{\mathbf{n}}'$.

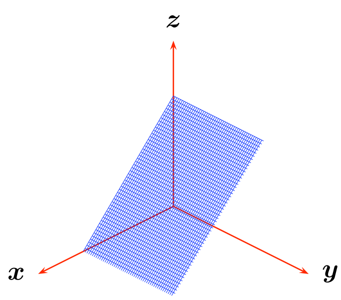

Figure 1.66: The plane $z = 1 - x$.

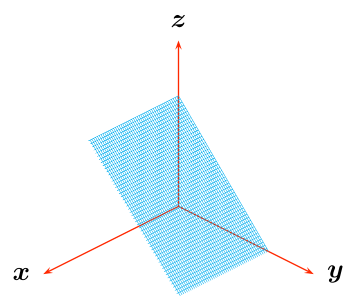

Figure 1.67: The plane $z = 1 - y$.

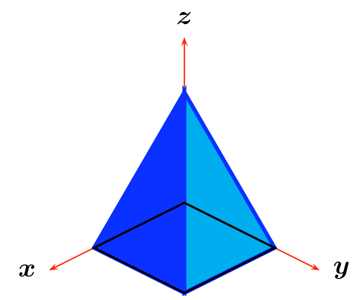

Figure 1.68: Solid bounded by the planes $z = 1 - x$ and $z = 1 - y$ in the first octant.

73 Example

1. Draw the intersection of the plane $z = 1 - x$ with the first octant.
2. Draw the intersection of the plane $z = 1 - y$ with the first octant.
3. Find the angle between the planes $z = 1 - x$ and $z = 1 - y$.
4. Draw the solid $\mathscr{S}$ which results from the intersection of the planes $z = 1 - x$ and $z = 1 - y$ with the first octant.
5. Find the volume of the solid $\mathscr{S}$.

Solution: ▶

1. *This appears in figure 1.66.*
2. *This appears in figure 1.67.*
3. *The vector* $\begin{bmatrix}1\\0\\1\end{bmatrix}$ *is normal to the plane* $x + z = 1$, *and the vector* $\begin{bmatrix}0\\1\\1\end{bmatrix}$ *is normal to the plane* $y + z = 1$. *If* θ *is the angle between these two vectors, then*
$$\cos\theta = \frac{1\cdot 0 + 0\cdot 1 + 1\cdot 1}{\sqrt{1^2+1^2}\cdot\sqrt{1^2+1^2}} \implies \cos\theta = \frac{1}{2} \implies \theta = \frac{\pi}{3}.$$
4. *This appears in figure 1.68.*
5. *The resulting solid is a pyramid with square base of area* $A = 1\cdot 1 = 1$. *Recall that the volume of a pyramid is given by the formula* $V = \frac{Ah}{3}$, *where* A *is area of the base of the pyramid and* h *is its height. Now, the height of this pyramid is clearly* 1, *and hence the volume required is* $\frac{1}{3}$.

◀

Homework

Problem 1.7.1 Vectors $\overrightarrow{a}$, $\overrightarrow{b}$ satisfy $\left|\left|\overrightarrow{a}\right|\right| = 13$, $\left|\left|\overrightarrow{b}\right|\right| = 19$, $\left|\left|\overrightarrow{a} + \overrightarrow{b}\right|\right| = 24$. Find $\left|\left|\overrightarrow{a} - \overrightarrow{b}\right|\right|$.

Problem 1.7.2 Find the equation of the line passing through $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ in the direction of $\begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix}$.

Problem 1.7.3 Find the equation of plane containing the point $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and perpendicular to the line $x = 1 + t, y = -2t, z = 1 - t$.

Problem 1.7.4 Find the equation of plane containing the point $\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$ and containing the line $x = 2y = 3z$.

Problem 1.7.5 (Putnam Exam 1984) Let A be a solid $a \times b \times c$ rectangular brick in three dimensions, where $a > 0, b > 0, c > 0$. Let B be the set of all points which are at distance at most 1 from some point of A (in particular, $A \subseteq B$). Express the volume of B as a polynomial in a, b, c.

Problem 1.7.6 It is known that $\left|\left|\overrightarrow{a}\right|\right| = 3$, $\left|\left|\overrightarrow{b}\right|\right| = 4$, $\left|\left|\overrightarrow{c}\right|\right| = 5$ and that $\overrightarrow{a} + \overrightarrow{b} + \overrightarrow{c} = \overrightarrow{0}$. Find

$$\overrightarrow{a}\bullet\overrightarrow{b} + \overrightarrow{b}\bullet\overrightarrow{c} + \overrightarrow{c}\bullet\overrightarrow{a}.$$

Problem 1.7.7 Find the equation of the line perpendicular to the plane $ax + a^2y + a^3z = 0, \quad a \neq 0$ and passing through the point $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Problem 1.7.8 Find the equation of the plane perpendicular to the line $ax = by = cz, \quad abc \neq 0$ and passing through the point $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ in $\mathbb{R}^3$.

Problem 1.7.9 Find the (shortest) distance from the point $(1, 2, 3)$ to the plane $x - y + z = 1$.

Problem 1.7.10 Determine whether the lines

$$L_1 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + t\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix},$$

$$L_2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + t\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix},$$

intersect. Find the angle between them.

Problem 1.7.11 Let a, b, c be arbitrary real numbers. Prove that

$$(a^2 + b^2 + c^2)^2 \leq 3(a^4 + b^4 + c^4).$$

Problem 1.7.12 Let $a > 0, b > 0, c > 0$ be the lengths of the sides of $\triangle ABC$. (Vertex A is opposite to the side measuring a, etc.) Recall that by Heron's Formula, the area of this triangle is $S(a, b, c) = \sqrt{s(s-a)(s-b)(s-c)}$, where $s = \dfrac{a+b+c}{2}$ is the semiperimeter of the triangle. Prove that $f(a, b, c) = \dfrac{S(a,b,c)}{a^2+b^2+c^2}$ is maximised when $\triangle ABC$ is equilateral, and find this maximum.

Problem 1.7.13 Let x, y, z be strictly positive numbers. Prove that

$$\frac{\sqrt{x+y} + \sqrt{y+z} + \sqrt{z+x}}{\sqrt{x+y+z}} \leq \sqrt{6}.$$

Problem 1.7.14 Let x, y, z be strictly positive numbers. Prove that

$$x + y + z \leq 2\left(\frac{x^2}{y+z} + \frac{y^2}{z+x} + \frac{z^2}{x+y}\right).$$

Problem 1.7.15 Find the Cartesian equation of the plane passing through $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Draw this plane and its intersection with the first octant. Find the volume of the tetrahedron with vertices at $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Problem 1.7.16 Prove that there do not exist three unit vectors in $\mathbb{R}^3$ such that the angle between any two of them be $> \dfrac{2\pi}{3}$.

Problem 1.7.17 Let $(\overrightarrow{r} - \overrightarrow{a})\bullet\overrightarrow{n} = 0$ be a plane passing through the point $\mathbf{a}$ and perpendicular to vector $\overrightarrow{n}$. If $\mathbf{b}$ is not a point on the plane, then the distance from $\mathbf{b}$ to the plane is

$$\frac{\left|(\overrightarrow{a} - \overrightarrow{b})\bullet\overrightarrow{n}\right|}{\left|\left|\overrightarrow{n}\right|\right|}.$$

Problem 1.7.18 (Putnam Exam 1980) Let S be the solid in three-dimensional space consisting of all points $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ satisfying the following system of six conditions:

$$x \geq 0, \quad y \geq 0, \quad z \geq 0,$$
$$x + y + z \leq 11,$$
$$2x + 4y + 3z \leq 36,$$
$$2x + 3z \leq 24.$$

Determine the number of vertices and the number of edges of S.

Problem 1.7.19 Given a polyhedron with n faces, consider n vectors, each normal to a face of the polyhedron, and length equal to the area of the face. Prove that the sum of these vectors is $\overrightarrow{0}$.

1.8 Cross Product

We now define the standard cross product in $\mathbb{R}^3$ as a product satisfying the following properties.

74 Definition Let $\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}$ be vectors in $\mathbb{R}^3$, and let $\alpha \in \mathbb{R}$ be a scalar. The cross product $\times$ is a closed binary operation satisfying

❶ **Anti-commutativity:** $\overrightarrow{x} \times \overrightarrow{y} = -(\overrightarrow{y} \times \overrightarrow{x})$

❷ **Bilinearity:**

$$(\overrightarrow{x} + \overrightarrow{z}) \times \overrightarrow{y} = \overrightarrow{x} \times \overrightarrow{y} + \overrightarrow{z} \times \overrightarrow{y} \quad \text{and} \quad \overrightarrow{x} \times (\overrightarrow{z} + \overrightarrow{y}) = \overrightarrow{x} \times \overrightarrow{z} + \overrightarrow{x} \times \overrightarrow{y}$$

❸ **Scalar homogeneity:** $(\alpha\overrightarrow{x}) \times \overrightarrow{y} = \overrightarrow{x} \times (\alpha\overrightarrow{y}) = \alpha(\overrightarrow{x} \times \overrightarrow{y})$

❹ $\overrightarrow{x} \times \overrightarrow{x} = \overrightarrow{0}$

❺ **Right-hand Rule:**

$$\overrightarrow{i} \times \overrightarrow{j} = \overrightarrow{k}, \quad \overrightarrow{j} \times \overrightarrow{k} = \overrightarrow{i}, \quad \overrightarrow{k} \times \overrightarrow{i} = \overrightarrow{j}.$$

It follows that the cross product is an operation that, given two non-parallel vectors on a plane, allows us to "get out" of that plane.

75 Example Find

$$\begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

Solution: ► *We have*

$$\begin{aligned} (\overrightarrow{i} - 3\overrightarrow{k}) \times (\overrightarrow{j} + 2\overrightarrow{k}) &= \overrightarrow{i} \times \overrightarrow{j} + 2\overrightarrow{i} \times \overrightarrow{k} - 3\overrightarrow{k} \times \overrightarrow{j} - 6\overrightarrow{k} \times \overrightarrow{k} \\ &= \overrightarrow{k} - 2\overrightarrow{j} + 3\overrightarrow{i} + 6\overrightarrow{0} \\ &= 3\overrightarrow{i} - 2\overrightarrow{j} + \overrightarrow{k}. \end{aligned}$$

Hence

$$\begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}.$$

◄

☞ *The cross product of vectors in $\mathbb{R}^3$ is not associative, since*

$$\overrightarrow{i} \times (\overrightarrow{i} \times \overrightarrow{j}) = \overrightarrow{i} \times \overrightarrow{k} = -\overrightarrow{j}$$

but

$$(\overrightarrow{i} \times \overrightarrow{i}) \times \overrightarrow{j} = \overrightarrow{0} \times \overrightarrow{j} = \overrightarrow{0}.$$

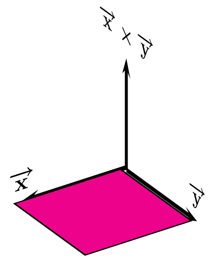

Figure 1.69: Theorem 79.

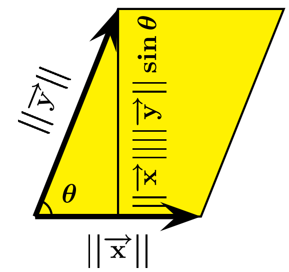

Figure 1.70: Area of a parallelogram

Operating as in example 75 we obtain

76 Theorem Let $\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\overrightarrow{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ be vectors in $\mathbb{R}^3$. Then

$$\overrightarrow{x} \times \overrightarrow{y} = (x_2y_3 - x_3y_2)\overrightarrow{i} + (x_3y_1 - x_1y_3)\overrightarrow{j} + (x_1y_2 - x_2y_1)\overrightarrow{k}.$$

Proof: *Since* $\overrightarrow{i} \times \overrightarrow{i} = \overrightarrow{j} \times \overrightarrow{j} = \overrightarrow{k} \times \overrightarrow{k} = \overrightarrow{0}$, *we only worry about the mixed products, obtaining,*

$$\begin{aligned}
\overrightarrow{x} \times \overrightarrow{y} &= (x_1\overrightarrow{i} + x_2\overrightarrow{j} + x_3\overrightarrow{k}) \times (y_1\overrightarrow{i} + y_2\overrightarrow{j} + y_3\overrightarrow{k}) \\
&= x_1y_2\overrightarrow{i} \times \overrightarrow{j} + x_1y_3\overrightarrow{i} \times \overrightarrow{k} + x_2y_1\overrightarrow{j} \times \overrightarrow{i} + x_2y_3\overrightarrow{j} \times \overrightarrow{k} \\
&\qquad + x_3y_1\overrightarrow{k} \times \overrightarrow{i} + x_3y_2\overrightarrow{k} \times \overrightarrow{j} \\
&= (x_1y_2 - y_1x_2)\overrightarrow{i} \times \overrightarrow{j} + (x_2y_3 - x_3y_2)\overrightarrow{j} \times \overrightarrow{k} + (x_3y_1 - x_1y_3)\overrightarrow{k} \times \overrightarrow{i} \\
&= (x_1y_2 - y_1x_2)\overrightarrow{k} + (x_2y_3 - x_3y_2)\overrightarrow{i} + (x_3y_1 - x_1y_3)\overrightarrow{j},
\end{aligned}$$

proving the theorem. ❑

Using the cross product, we may obtain a third vector simultaneously perpendicular to two other vectors in space.

77 Theorem $\overrightarrow{x} \perp (\overrightarrow{x} \times \overrightarrow{y})$ and $\overrightarrow{y} \perp (\overrightarrow{x} \times \overrightarrow{y})$, that is, the cross product of two vectors is simultaneously perpendicular to both original vectors.

Proof: *We will only check the first assertion, the second verification is analogous.*

$$\begin{aligned}
\overrightarrow{x} \bullet (\overrightarrow{x} \times \overrightarrow{y}) &= (x_1\overrightarrow{i} + x_2\overrightarrow{j} + x_3\overrightarrow{k}) \bullet ((x_2y_3 - x_3y_2)\overrightarrow{i} \\
&\qquad + (x_3y_1 - x_1y_3)\overrightarrow{j} + (x_1y_2 - x_2y_1)\overrightarrow{k}) \\
&= x_1x_2y_3 - x_1x_3y_2 + x_2x_3y_1 - x_2x_1y_3 + x_3x_1y_2 - x_3x_2y_1 \\
&= 0,
\end{aligned}$$

completing the proof. ❑

Although the cross product is not associative, we have, however, the following theorem.

78 Theorem

$$\overrightarrow{a} \times (\overrightarrow{b} \times \overrightarrow{c}) = (\overrightarrow{a} \bullet \overrightarrow{c})\overrightarrow{b} - (\overrightarrow{a} \bullet \overrightarrow{b})\overrightarrow{c}.$$

Proof:

$$\begin{aligned}\overrightarrow{a}\times(\overrightarrow{b}\times\overrightarrow{c}) &= (a_1\overrightarrow{i}+a_2\overrightarrow{j}+a_3\overrightarrow{k})\times((b_2c_3-b_3c_2)\overrightarrow{i}+\\ &\quad +(b_3c_1-b_1c_3)\overrightarrow{j}+(b_1c_2-b_2c_1)\overrightarrow{k})\\ &= a_1(b_3c_1-b_1c_3)\overrightarrow{k}-a_1(b_1c_2-b_2c_1)\overrightarrow{j}-a_2(b_2c_3-b_3c_2)\overrightarrow{k}\\ &\quad +a_2(b_1c_2-b_2c_1)\overrightarrow{i}+a_3(b_2c_3-b_3c_2)\overrightarrow{j}-a_3(b_3c_1-b_1c_3)\overrightarrow{i}\\ &= (a_1c_1+a_2c_2+a_3c_3)(b_1\overrightarrow{i}+b_2\overrightarrow{j}+b_3\overrightarrow{i})+\\ &\quad (-a_1b_1-a_2b_2-a_3b_3)(c_1\overrightarrow{i}+c_2\overrightarrow{j}+c_3\overrightarrow{i})\\ &= (\overrightarrow{a}\bullet\overrightarrow{c})\overrightarrow{b}-(\overrightarrow{a}\bullet\overrightarrow{b})\overrightarrow{c},\end{aligned}$$

completing the proof. ❑

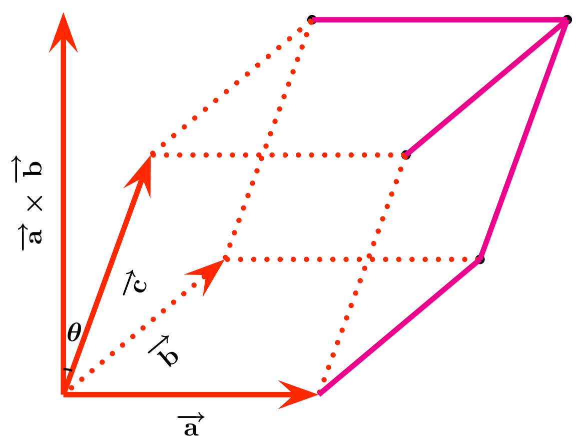

Figure 1.71: Theorem 82.

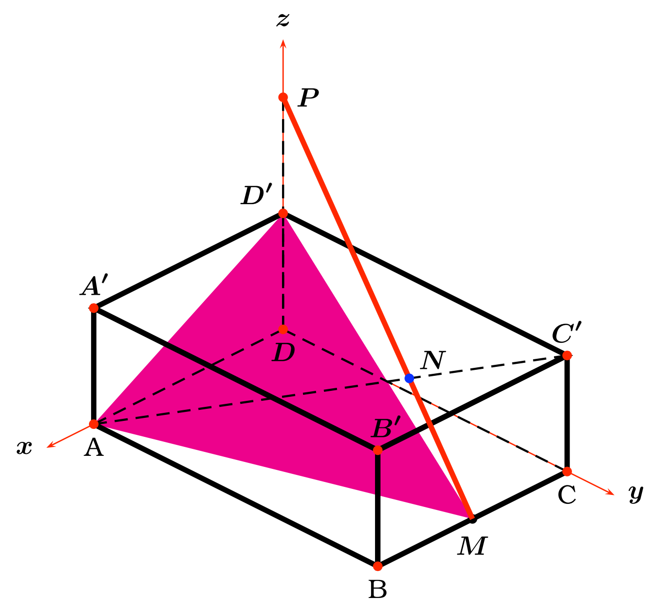

Figure 1.72: Example 83.

79 Theorem Let $(\widehat{\overrightarrow{x},\overrightarrow{y}})\in[0;\pi]$ be the convex angle between two vectors $\overrightarrow{x}$ and $\overrightarrow{y}$. Then

$$||\overrightarrow{x}\times\overrightarrow{y}|| = ||\overrightarrow{x}||||\overrightarrow{y}||\sin(\widehat{\overrightarrow{x},\overrightarrow{y}}).$$

Proof: *We have*

$$\begin{aligned}||\overrightarrow{x}\times\overrightarrow{y}||^2 &= (x_2y_3-x_3y_2)^2+(x_3y_1-x_1y_3)^2+(x_1y_2-x_2y_1)^2\\ &= y^2y_3^2-2x_2y_3x_3y_2+z^2y_2^2+z^2y_1^2-2x_3y_1x_1y_3+\\ &\quad +x^2y_3^2+x^2y_2^2-2x_1y_2x_2y_1+y^2y_1^2\\ &= (x^2+y^2+z^2)(y_1^2+y_2^2+y_3^2)-(x_1y_1+x_2y_2+x_3y_3)^2\\ &= ||\overrightarrow{x}||^2||\overrightarrow{y}||^2-(\overrightarrow{x}\bullet\overrightarrow{y})^2\\ &= ||\overrightarrow{x}||^2||\overrightarrow{y}||^2-||\overrightarrow{x}||^2||\overrightarrow{y}||^2\cos^2(\widehat{\overrightarrow{x},\overrightarrow{y}})\\ &= ||\overrightarrow{x}||^2||\overrightarrow{y}||^2\sin^2(\widehat{\overrightarrow{x},\overrightarrow{y}}),\end{aligned}$$

whence the theorem follows. ❑

Theorem 79 has the following geometric significance: $\left|\left|\overrightarrow{x} \times \overrightarrow{y}\right|\right|$ is the area of the parallelogram formed when the tails of the vectors are joined. See figure 1.70.
The following corollaries easily follow from Theorem 79.

80 Corollary Two non-zero vectors $\overrightarrow{x}$, $\overrightarrow{y}$ satisfy $\overrightarrow{x} \times \overrightarrow{y} = \overrightarrow{0}$ if and only if they are parallel.

81 Corollary (Lagrange's Identity)

$$\left|\left|\overrightarrow{x} \times \overrightarrow{y}\right|\right|^2 = ||x||^2||y||^2 - (\overrightarrow{x}\bullet\overrightarrow{y})^2.$$

The following result mixes the dot and the cross product.

82 Theorem Let $\overrightarrow{a}$, $\overrightarrow{b}$, $\overrightarrow{c}$, be linearly independent vectors in $\mathbb{R}^3$. The signed volume of the parallelepiped spanned by them is $(\overrightarrow{a} \times \overrightarrow{b}) \bullet \overrightarrow{c}$.

Proof: *See figure 1.71. The area of the base of the parallelepiped is the area of the parallelogram determined by the vectors $\overrightarrow{a}$ and $\overrightarrow{b}$, which has area $\left|\left|\overrightarrow{a} \times \overrightarrow{b}\right|\right|$. The altitude of the parallelepiped is $\left|\left|\overrightarrow{c}\right|\right| \cos\theta$ where θ is the angle between $\overrightarrow{c}$ and $\overrightarrow{a} \times \overrightarrow{b}$. The volume of the parallelepiped is thus*

$$\left|\left|\overrightarrow{a} \times \overrightarrow{b}\right|\right|\left|\left|\overrightarrow{c}\right|\right| \cos\theta = (\overrightarrow{a} \times \overrightarrow{b}) \bullet \overrightarrow{c},$$

proving the theorem. ❑

☞ *Since we may have used any of the faces of the parallelepiped, it follows that*

$$(\overrightarrow{a} \times \overrightarrow{b}) \bullet \overrightarrow{c} = (\overrightarrow{b} \times \overrightarrow{c}) \bullet \overrightarrow{a} = (\overrightarrow{c} \times \overrightarrow{a}) \bullet \overrightarrow{b}.$$

In particular, it is possible to "exchange" the cross and dot products:

$$\overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c}) = (\overrightarrow{a} \times \overrightarrow{b}) \bullet \overrightarrow{c}$$

83 Example Consider the rectangular parallelepiped $ABCDD'C'B'A'$ (figure 1.72) with vertices $A(2,0,0)$, $B(2,3,0)$, $C(0,3,0)$, $D(0,0,0)$, $D'(0,0,1)$, $C'(0,3,1)$, $B'(2,3,1)$, $A'(2,0,1)$. Let M be the midpoint of the line segment joining the vertices B and C.

1. Find the Cartesian equation of the plane containing the points A, D', and M.
2. Find the area of $\triangle AD'M$.
3. Find the parametric equation of the line $\overleftrightarrow{AC'}$.
4. Suppose that a line through M is drawn cutting the line segment $[AC']$ in N and the line $\overleftrightarrow{DD'}$ in P. Find the parametric equation of $\overleftrightarrow{MP}$.

Solution: ►

1. *Form the following vectors and find their cross product:*

$$\overrightarrow{AD'} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \qquad \overrightarrow{AM} = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} \implies \overrightarrow{AD'} \times \overrightarrow{AM} = \begin{bmatrix} -3 \\ -1 \\ -6 \end{bmatrix}.$$

The equation of the plane is thus

$$\begin{bmatrix} x-2 \\ y-0 \\ z-0 \end{bmatrix} \bullet \begin{bmatrix} -3 \\ -1 \\ -6 \end{bmatrix} = 0 \implies 3(x-2) + 1(y) + 6z = 0 \implies 3x + y + 6z = 6.$$

2. *The area of the triangle is*

$$\frac{\left\|\overrightarrow{\mathbf{AD'}} \times \overrightarrow{\mathbf{AM}}\right\|}{2} = \frac{1}{2}\sqrt{3^2+1^2+6^2} = \frac{\sqrt{46}}{2}.$$

3. *We have* $\overrightarrow{\mathbf{AC'}} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$, *and hence the line* $\overleftrightarrow{\mathbf{AC'}}$ *has parametric equation*

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + t\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix} \implies x = 2-2t,\ y = 3t,\ z = t.$$

4. *Since P is on the z-axis,* $P = \begin{pmatrix} 0 \\ 0 \\ z' \end{pmatrix}$ *for some real number* $z' > 0$. *The parametric equation of the line* $\overleftrightarrow{\mathbf{MP}}$ *is thus*

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + s\begin{bmatrix} -1 \\ -3 \\ z' \end{bmatrix} \implies x = 1-s \qquad y = 3-3s, \qquad z = sz'.$$

Since N is on both $\overleftrightarrow{\mathbf{MP}}$ *and* $\overleftrightarrow{\mathbf{AC'}}$ *we must have*

$$2-2t = 1-s,\ 3t = 3-3s,\ t = sz'.$$

Solving the first two equations gives $s = \frac{1}{3}, t = \frac{2}{3}$. *Putting this into the third equation we deduce* $z' = 2$. *Thus* $P = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$ *and the desired equation is*

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + s\begin{bmatrix} -1 \\ -3 \\ 2 \end{bmatrix} \implies x = 1-s, \quad y = 3-3s, \qquad z = 2s.$$

◀

Homework

Problem 1.8.1 Prove that

$$(\overrightarrow{a} - \overrightarrow{b}) \times (\overrightarrow{a} + \overrightarrow{b}) = 2\overrightarrow{a} \times \overrightarrow{b}.$$

Problem 1.8.2 Prove that $\overrightarrow{x} \times \overrightarrow{x} = \overrightarrow{0}$ follows from the anti-commutativity of the cross product.

Problem 1.8.3 If $\overrightarrow{b} - \overrightarrow{a}$ and $\overrightarrow{c} - \overrightarrow{a}$ are parallel and it is known that $\overrightarrow{c} \times \overrightarrow{a} = \overrightarrow{i} - \overrightarrow{j}$ and $\overrightarrow{a} \times \overrightarrow{b} = \overrightarrow{j} + \overrightarrow{k}$, find $\overrightarrow{b} \times \overrightarrow{c}$.

Problem 1.8.4 Redo example 71, that is, find the Cartesian equation of the plane parallel to the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and passing through the point $(0, -1, 2)$, by finding a normal to the plane.

Problem 1.8.5 Find the equation of the plane passing through the points $(a, 0, a)$, $(-a, 1, 0)$, and $(0, 1, 2a)$ in $\mathbb{R}^3$.

Problem 1.8.6 Let $a \in \mathbb{R}$. Find a vector of unit length simultaneously perpendicular to $\overrightarrow{v} = \begin{bmatrix} 0 \\ -a \\ a \end{bmatrix}$ and $\overrightarrow{w} = \begin{bmatrix} 1 \\ a \\ 0 \end{bmatrix}$.

Problem 1.8.7 (Jacobi's Identity) Let $\overrightarrow{a}$, $\overrightarrow{b}$, $\overrightarrow{c}$ be vectors in $\mathbb{R}^3$. Prove that

$$\overrightarrow{a} \times (\overrightarrow{b} \times \overrightarrow{c}) + \overrightarrow{b} \times (\overrightarrow{c} \times \overrightarrow{a}) + \overrightarrow{c} \times (\overrightarrow{a} \times \overrightarrow{b}) = \overrightarrow{0}.$$

Problem 1.8.8 Let $\overrightarrow{x} \in \mathbb{R}^3, ||\mathrm{x}|| = 1$. Find

$$||\overrightarrow{x} \times \overrightarrow{i}||^2 + ||\overrightarrow{x} \times \overrightarrow{j}||^2 + ||\overrightarrow{x} \times \overrightarrow{k}||^2.$$

Problem 1.8.9 The vectors $\overrightarrow{a}$, $\overrightarrow{b}$ are constant vectors. Solve the equation

$$\overrightarrow{a} \times (\overrightarrow{x} \times \overrightarrow{b}) = \overrightarrow{b} \times (\overrightarrow{x} \times \overrightarrow{a}).$$

Problem 1.8.10 If $\overrightarrow{a} + \overrightarrow{b} + \overrightarrow{c} = \overrightarrow{0}$, prove that

$$\overrightarrow{a} \times \overrightarrow{b} = \overrightarrow{b} \times \overrightarrow{c} = \overrightarrow{c} \times \overrightarrow{a}.$$

Problem 1.8.11 Assume $\overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c}) \neq 0$ and that

$$\overrightarrow{x} = \alpha\overrightarrow{a} + \beta\overrightarrow{b} + \gamma\overrightarrow{c}.$$

Find α, β, and γ in terms of $\overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c})$.

Problem 1.8.12 The vectors $\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}$ are constant vectors. Solve the system of equations

$$2\overrightarrow{x} + \overrightarrow{y} \times \overrightarrow{a} = \overrightarrow{b}, \quad 3\overrightarrow{y} + \overrightarrow{x} \times \overrightarrow{a} = \overrightarrow{c},$$

Problem 1.8.13 Let $\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}, \overrightarrow{d}$ be vectors in $\mathbb{R}^3$. Prove the following vector identity,

$$(\overrightarrow{a} \times \overrightarrow{b}) \bullet (\overrightarrow{c} \times \overrightarrow{d}) = (\overrightarrow{a} \bullet \overrightarrow{c})(\overrightarrow{b} \bullet \overrightarrow{d}) - (\overrightarrow{a} \bullet \overrightarrow{d})(\overrightarrow{b} \bullet \overrightarrow{c}).$$

Problem 1.8.14 Let $\overrightarrow{a}$, $\overrightarrow{b}$, $\overrightarrow{c}$, $\overrightarrow{d}$, be vectors in $\mathbb{R}^3$. Prove that

$$\begin{aligned} (\overrightarrow{b} \times \overrightarrow{c}) \bullet (\overrightarrow{a} \times \overrightarrow{d}) & \\ & +(\overrightarrow{c} \times \overrightarrow{a}) \bullet (\overrightarrow{b} \times \overrightarrow{d}) \\ & +(\overrightarrow{a} \times \overrightarrow{b}) \bullet (\overrightarrow{c} \times \overrightarrow{d}) \\ & = 0. \end{aligned}$$

Problem 1.8.15 Consider the plane Π passing through the points $A(6, 0, 0)$, $B(0, 4, 0)$ and $C(0, 0, 3)$, as shewn in figure 1.73 below. The plane Π intersects a $3 \times 3 \times 3$ cube, one of whose vertices is at the origin and that has three of its edges on the coordinate axes, as in the figure. This intersection forms a pentagon $CPQRS$.

1. Find $\overrightarrow{\mathbf{CA}} \times \overrightarrow{\mathbf{CB}}$.
2. Find $\left|\left|\overrightarrow{\mathbf{CA}} \times \overrightarrow{\mathbf{CB}}\right|\right|$.
3. Find the parametric equation of the line L_{CA} joining C and A, with a parameter $t \in \mathbb{R}$.
4. Find the parametric equation of the line L_{DE} joining D and E, with a parameter $s \in \mathbb{R}$.
5. Find the intersection point between the lines L_{CA} and L_{DE}.
6. Find the equation of the plane Π.
7. Find the area of $\triangle ABC$.
8. Find the coordinates of the points P, Q, R, and S.
9. Find the area of the pentagon $CPQRS$.

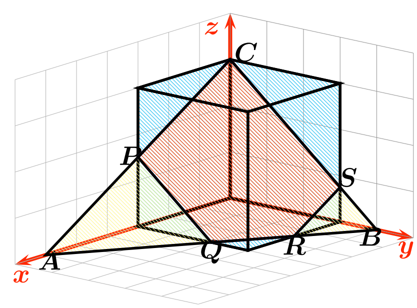

Figure 1.73: Problem 1.8.15..

1.9 Matrices in three dimensions

We will briefly introduce 3×3 matrices. Most of the material will flow like that for 2×2 matrices.

84 Definition A *linear transformation* $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a function such that

$$T(\mathrm{a} + \mathrm{b}) = T(\mathrm{a}) + T(\mathrm{b}), \qquad T(\lambda\mathrm{a}) = \lambda T(\mathrm{a}),$$

for all points a, b in $\mathbb{R}^3$ and all scalars λ. Such a linear transformation has a 3×3 matrix representation whose columns are the vectors $T(\mathrm{i})$, $T(\mathrm{j})$, and $T(\mathrm{k})$.

85 Example Consider $L : \mathbb{R}^3 \to \mathbb{R}^3$, with

$$L\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 - x_2 - x_3 \\ x_1 + x_2 + x_3 \\ x_3 \end{pmatrix}.$$

❶ Prove that L is a linear transformation.

❷ Find the matrix corresponding to L under the standard basis.

Solution: ►

❶ *Let $\alpha \in \mathbb{R}$ and let $\mathbf{u}, \mathbf{v}$ be points in $\mathbb{R}^3$. Then*

$$\begin{aligned} L(\mathbf{u}+\mathbf{v}) &= L\left(\begin{pmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{pmatrix}\right) \\ &= \begin{pmatrix} (u_1+v_1)-(u_2+v_2)-(u_3+v_3) \\ (u_1+v_1)+(u_2+v_2)+(u_3+v_3) \\ u_3+v_3 \end{pmatrix} \\ &= \begin{pmatrix} u_1-u_2-u_3 \\ u_1+u_2+u_3 \\ u_3 \end{pmatrix} + \begin{pmatrix} v_1-v_2-v_3 \\ v_1+v_2+v_3 \\ v_3 \end{pmatrix} \\ &= L\left(\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}\right) + L\left(\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}\right) \\ &= L(\mathbf{u}) + L(\mathbf{v}), \end{aligned}$$

and also

$$\begin{aligned} L(\alpha\mathbf{u}) &= L\left(\begin{pmatrix} \alpha u_1 \\ \alpha u_2 \\ \alpha u_3 \end{pmatrix}\right) \\ &= \begin{pmatrix} \alpha(u_1)-\alpha(u_2)-\alpha(u_3) \\ \alpha(u_1)+\alpha(u_2)+\alpha(u_3) \\ \alpha u_3 \end{pmatrix} \\ &= \alpha\begin{pmatrix} u_1-u_2-u_3 \\ u_1+u_2+u_3 \\ u_3 \end{pmatrix} \\ &= \alpha L\left(\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}\right) \\ &= \alpha L(\mathbf{u}), \end{aligned}$$

proving that L is a linear transformation.

❷ *We have* $L\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $L\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, *and* $L\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, *whence the*

desired matrix is

$$\begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

◀

Addition, scalar multiplication, and matrix multiplication are defined for 3×3 matrices in a manner analogous to those operations for 2×2 matrices.

86 Definition Let A, B be 3×3 matrices. Then we define

$$A + B = [a_{ij} + b_{ij}], \qquad \alpha A = [\alpha a_{ij}], \qquad AB = \left[\sum_{k=1}^{3} a_{ik} b_{kj}\right].$$

87 Example If $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \\ 6 & 0 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} a & b & c \\ a & b & 0 \\ a & 0 & 0 \end{bmatrix}$, then

$$A + B = \begin{bmatrix} 1+a & 2+b & 3+c \\ 4+a & 5+b & 0 \\ 6+a & 0 & 0 \end{bmatrix}, \qquad 3A = \begin{bmatrix} 3 & 6 & 9 \\ 12 & 15 & 0 \\ 18 & 0 & 0 \end{bmatrix},$$

$$BA = \begin{bmatrix} a+4b+6c & 2a+5b & 3a \\ a+4b & 2a+5b & 3a \\ a & 2a & 3a \end{bmatrix}, \qquad AB = \begin{bmatrix} 6a & 3b & c \\ 9a & 9b & 4c \\ 6a & 6b & 6c \end{bmatrix}.$$

88 Definition A *scaling matrix* is one of the form

$$S_{a,b,c} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix},$$

where $a > 0$, $b > 0$, $c > 0$.

It is an easy exercise to prove that the product of two scaling matrices commutes.

89 Definition A *rotation matrix* about the z-axis by an angle θ in the counterclockwise sense is

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A *rotation matrix* about the y-axis by an angle θ in the counterclockwise sense is

$$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}.$$

A *rotation matrix* about the x-axis by an angle θ in the counterclockwise sense is

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

Easy to find counterexamples should convince the reader that the product of two rotations in space does not necessarily commute.

90 Definition A *reflexion matrix* about the x-axis is

$$R_x = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A *reflexion matrix* about the y-axis is

$$R_y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A *reflexion matrix* about the z-axis is

$$R_z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Homework

Problem 1.9.1 Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. Find A^2, A^3 and A^4. Conjecture and, prove by induction, a general formula for A^n.

Problem 1.9.2 Let $A \in \mathrm{M}_{3\times 3}(\mathbb{R})$ be given by

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Demonstrate, using induction, that $A^n = 3^{n-1}A$ for $n \in \mathbb{N}, n \geq 1$.

Problem 1.9.3 Consider the $n \times n$ matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & \ldots & 1 & 1 \\ 0 & 1 & 1 & 1 & \ldots & 1 & 1 \\ 0 & 0 & 1 & 1 & \ldots & 1 & 1 \\ \ldots & \ldots & \vdots & \vdots & \vdots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 \end{bmatrix}.$$

Describe A^2 and A^3 in terms of n.

Problem 1.9.4 Let x be a real number, and put

$$m(x) = \begin{bmatrix} 1 & 0 & x \\ -x & 1 & -\frac{x^2}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

If a, b are real numbers, prove that

1. $m(a)m(b) = m(a+b)$.
2. $m(a)m(-a) = \mathbf{I}_3$, the 3×3 identity matrix.

1.10 Determinants in three dimensions

We now define the notion of *determinant* of a 3×3 matrix. Consider now the vectors $\overrightarrow{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$, $\overrightarrow{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, $\overrightarrow{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$, in $\mathbb{R}^3$, and the 3×3 matrix $A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$. Since thanks to Theorem 82,

the volume of the parallelepiped spanned by these vectors is $\overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c})$, we *define* the determinant of A, $\det A$, to be

$$D(\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}) = \det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = \overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c}). \tag{1.15}$$

We now establish that the properties of the determinant of a 3×3 as defined above are analogous to those of the determinant of 2×2 matrix defined in the preceding chapter.

91 Theorem The determinant of a 3×3 matrix A as defined by (1.15) satisfies the following properties:

1. D is linear in each of its arguments.
2. If the parallelepiped is flat then the volume is 0, that is, if $\overrightarrow{a}$, $\overrightarrow{b}$, $\overrightarrow{c}$, are linearly dependent, then $D(\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}) = 0$.
3. $D(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}) = 1$, and accords with the right-hand rule.

Proof:

1. *If* $D(\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}) = \overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c})$, *linearity of the first component follows by the distributive law for the dot product:*

$$\begin{aligned} D(\overrightarrow{a} + \overrightarrow{a}', \overrightarrow{b}, \overrightarrow{c}) &= (\overrightarrow{a} + \overrightarrow{a}') \bullet (\overrightarrow{b} \times \overrightarrow{c}) \\ &= \overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c}) + \overrightarrow{a}' \bullet (\overrightarrow{b} \times \overrightarrow{c}) \\ &= D(\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}) + D(\overrightarrow{a}', \overrightarrow{b}, \overrightarrow{c}), \end{aligned}$$

and if $\lambda \in \mathbb{R}$,

$$D(\lambda \overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}) = (\lambda \overrightarrow{a}) \bullet (\overrightarrow{b} \times \overrightarrow{c}) = \lambda((\overrightarrow{a}) \bullet (\overrightarrow{b} \times \overrightarrow{c})) = \lambda D(\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}).$$

The linearity on the second and third component can be established by using the distributive law of the cross product. For example, for the second component we have,

$$\begin{aligned} D(\overrightarrow{a}, \overrightarrow{b} + \overrightarrow{b}', \overrightarrow{c}) &= \overrightarrow{a} \bullet ((\overrightarrow{b} + \overrightarrow{b}') \times \overrightarrow{c}) \\ &= \overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c} + \overrightarrow{b'} \times \overrightarrow{c}) \\ &= \overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c}) + \overrightarrow{a} \bullet (\overrightarrow{b'} \times \overrightarrow{c}) \\ &= D(\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}) + D(\overrightarrow{a}, \overrightarrow{b'}, \overrightarrow{c}), \end{aligned}$$

and if $\lambda \in \mathbb{R}$,

$$D(\overrightarrow{a}, \lambda \overrightarrow{b}, \overrightarrow{c}) = \overrightarrow{a} \bullet ((\lambda \overrightarrow{b}) \times \overrightarrow{c}) = \lambda(\overrightarrow{a} \bullet (\overrightarrow{b} \times \overrightarrow{c})) = \lambda D(\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}).$$

2. *If* $\overrightarrow{a}$, $\overrightarrow{b}$, $\overrightarrow{c}$, *are linearly dependent, then they lie on the same plane and the parallelepiped spanned by them is flat, hence,* $D(\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}) = 0$.
3. *Since* $\overrightarrow{j} \times \overrightarrow{k} = \overrightarrow{i}$, *and* $\overrightarrow{i} \cdot \overrightarrow{i} = 1$,

$$D(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}) = \overrightarrow{i} \bullet (\overrightarrow{j} \times \overrightarrow{k}) = \overrightarrow{i} \cdot \overrightarrow{i} = 1.$$

❑

Observe that

$$\begin{aligned}
\det\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} &= \vec{a} \bullet (\vec{b} \times \vec{c}) && (1.16) \\
&= \vec{a} \bullet \left((b_2c_3 - b_3c_2)\vec{i} + (b_3c_1 - b_1c_3)\vec{j} + (b_1c_2 - b_2c_1)\vec{k}\right) && (1.17) \\
&= a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1) && (1.18) \\
&= a_1 \det\begin{bmatrix} b_2 & c_2 \\ b_3 & c_3 \end{bmatrix} - a_2 \det\begin{bmatrix} b_1 & c_1 \\ b_3 & c_3 \end{bmatrix} + a_3 \det\begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix}, && (1.19)
\end{aligned}$$

which reduces the computation of 3×3 determinants to 2×2 determinants.

92 Example Find $\det\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

Solution: ► *Using (1.19), we have*

$$\begin{aligned}
\det A &= 1\det\begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} - 4\det\begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix} + 7\det\begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix} \\
&= 1(45 - 48) - 4(18 - 24) + 7(12 - 15) \\
&= -3 + 24 - 21 \\
&= 0.
\end{aligned}$$

◄

Again, we may use the Maple™ packages `linalg`, `LinearAlgebra`, or `Student[VectorCalculus]` to perform many of the vector operations. An example follows with `linalg`.

```
> with(linalg):
> a:=vector([-2,0,1]);
```

$$a := [-2, 0, 1]$$

```
> b:=vector([-1,3,0]);
```

$$b := [-1, 3, 0]$$

```
> crossprod(a,b);
```

$$[-3, -1, -6]$$

```
> dotprod(a,b);
```

$$2$$

```
> angle(a,b);
```

$$\arccos\left(\frac{\sqrt{50}}{25}\right)$$

Homework

Problem 1.10.1 Prove that

$$\det\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix} = (b-c)(c-a)(a-b).$$

Problem 1.10.2 Prove that

$$\det\begin{bmatrix} a & b & c \\ b & c & a \\ c & a & b \end{bmatrix} = 3abc - a^3 - b^3 - c^3.$$

1.11 Some Solid Geometry

In this section we examine some examples and prove some theorems of three-dimensional geometry.

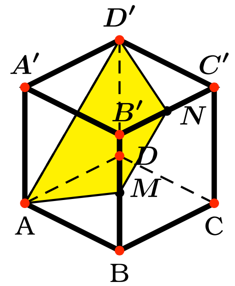

Figure 1.74: Example 93.

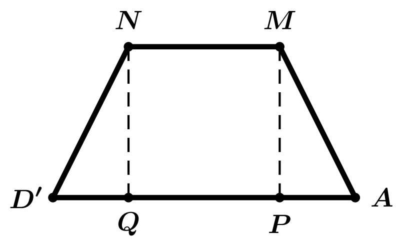

Figure 1.75: Example 93.

93 Example Cube $\mathrm{ABCDD'C'B'A}$ in figure 1.74 has side of length a. M is the midpoint of edge $[\mathrm{BB'}]$ and N is the midpoint of edge $[\mathrm{B'C'}]$. Prove that $\overrightarrow{\mathrm{AD'}} \parallel \overrightarrow{\mathrm{MN}}$ and find the area of the quadrilateral $\mathrm{MND'A}$.

Solution: ► *By the Pythagorean Theorem, $\left\|\overrightarrow{\mathrm{AD'}}\right\| = a\sqrt{2}$. Because they are diagonals that belong to parallel faces of the cube, $\overrightarrow{\mathrm{AD'}} \parallel \overrightarrow{\mathrm{BC'}}$. Now, M and N are the midpoints of the sides $[B'B]$ and $[B'C']$ of $\triangle B'C'B$, and hence $\overrightarrow{\mathrm{MN}} \parallel \overrightarrow{\mathrm{BC'}}$ by example 14. The aforementioned example also gives $\left\|\overrightarrow{\mathrm{MN}}\right\| = \frac{1}{2}\left\|\overrightarrow{\mathrm{AD'}}\right\| = \frac{a\sqrt{2}}{2}$. In consequence, $\overrightarrow{\mathrm{AD'}} \parallel \overrightarrow{\mathrm{MN}}$. This means that the four points $\mathrm{A}, \mathrm{D'}, \mathrm{M}, \mathrm{N}$ are all on the same plane. Hence $\mathrm{MND'A}$ is a trapezoid with bases of length $a\sqrt{2}$ and $\frac{a\sqrt{2}}{2}$ (figure 1.75). From the figure*

$$\left\|\overrightarrow{\mathrm{D'Q}}\right\| = \left\|\overrightarrow{\mathrm{AP}}\right\| = \frac{1}{2}\left(\left\|\overrightarrow{\mathrm{AD'}}\right\| - \left\|\overrightarrow{\mathrm{MN}}\right\|\right) = \frac{a\sqrt{2}}{4}.$$

Also, by the Pythagorean Theorem,

$$\left\|\overrightarrow{\mathrm{D'N}}\right\| = \sqrt{\left\|\overrightarrow{\mathrm{D'C'}}\right\|^2 + \left\|\overrightarrow{\mathrm{C'N}}\right\|^2} = \sqrt{a^2 + \frac{a^2}{4}} = \frac{a\sqrt{5}}{2}.$$

The height of this trapezoid is thus

$$\left\|\overrightarrow{\mathrm{NQ}}\right\| = \sqrt{\frac{5a^2}{4} - \frac{a^2}{8}} = \frac{3a}{2\sqrt{2}}.$$

The area of the trapezoid is finally,

$$\frac{3a}{2\sqrt{2}} \cdot \left(\frac{a\sqrt{2} + \frac{a\sqrt{2}}{2}}{2}\right) = \frac{9a^2}{8}.$$

◄

Let us prove a three-dimensional version of Thales' Theorem.

94 Theorem (Thales' Theorem) Of two lines are cut by three parallel planes, their corresponding segments are proportional.

Proof: *See figure 1.76. Given the lines* $\overleftrightarrow{\mathrm{AB}}$ *and* $\overleftrightarrow{\mathrm{CD}}$*, we must prove that*

$$\frac{\overline{AE}}{\overline{EB}} = \frac{\overline{CF}}{\overline{FD}}.$$

Draw line $\overleftrightarrow{\mathrm{AD}}$ *cutting plane* $P2$ *in* G*. The plane containing points* A, B, *and* D *intersects plane* $P2$ *in the line* $\overleftrightarrow{\mathrm{EG}}$*. Similarly the plane containing points* A, C, *and* D *intersects plane* $P2$ *in the line* $\overleftrightarrow{\mathrm{GF}}$*. Since* $P2$ *and* $P3$ *are parallel planes,* $\overleftrightarrow{\mathrm{EG}} \parallel \overleftrightarrow{\mathrm{BD}}$*, and so by Thales' Theorem on the plane (theorem 30),*

$$\frac{\overline{AE}}{\overline{EB}} = \frac{\overline{AG}}{\overline{GD}}.$$

Similarly, since $P1$ *and* $P2$ *are parallel,* $\overleftrightarrow{\mathrm{AC}} \parallel \overleftrightarrow{\mathrm{GF}}$ *and*

$$\frac{\overline{CF}}{\overline{FD}} = \frac{\overline{AG}}{\overline{GD}}.$$

It follows that

$$\frac{\overline{AE}}{\overline{EB}} = \frac{\overline{CF}}{\overline{FD}},$$

as needed to be shewn. ❑

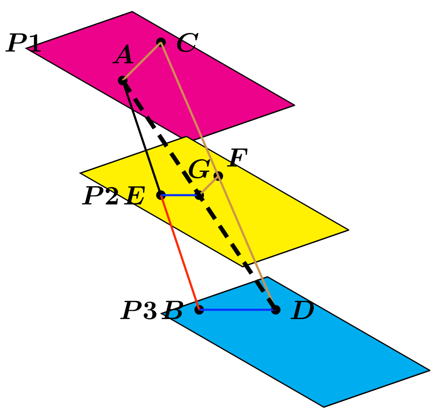

Figure 1.76: Thales' Theorem in 3D.

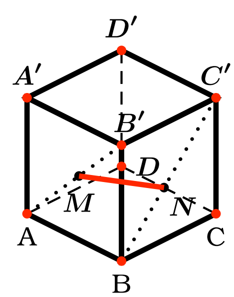

Figure 1.77: Example 95.

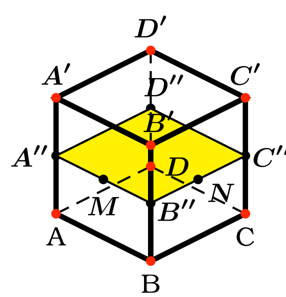

Figure 1.78: Example 95.

95 Example In cube $\mathrm{ABCDD'C'B'A'}$ of edge of length a, the points M and N are located on diagonals $[\mathrm{AB'}]$ and $[\mathrm{BC'}]$ such that $\overrightarrow{\mathrm{MN}}$ is parallel to the face ABCD of the cube. If $\left|\left|\overrightarrow{\mathrm{MN}}\right|\right| = \frac{\sqrt{5}}{3}\left|\left|\overrightarrow{\mathrm{AB}}\right|\right|$, find the ratios $\frac{\left|\left|\overrightarrow{\mathrm{AM}}\right|\right|}{\left|\left|\overrightarrow{\mathrm{AB'}}\right|\right|}$ and $\frac{\left|\left|\overrightarrow{\mathrm{BN}}\right|\right|}{\left|\left|\overrightarrow{\mathrm{BC'}}\right|\right|}$.

Solution: ▶ *There is a unique plane parallel* P *to face* ABCD *and containing* M*. Since* $\overrightarrow{\mathrm{MN}}$ *is parallel to face* ABCD *,* P *also contains* N*. The intersection of* P *with the cube produces a lamina* $\mathrm{A''B''C''D''}$*, as in figure 1.78.*

First notice that $\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right| = \left|\left|\overrightarrow{\mathbf{BC'}}\right|\right| = a\sqrt{2}$. *Put*

$$\frac{\left|\left|\overrightarrow{\mathbf{AM}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right|} = x \implies \frac{\left|\left|\overrightarrow{\mathbf{MB'}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right|} = \frac{\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right| - \left|\left|\overrightarrow{\mathbf{AM}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right|} = 1 - x.$$

Now, as $\triangle B'AB \sim \triangle B'MB''$ *and* $\triangle BC'B' \sim \triangle BNB''$,

$$\frac{\left|\left|\overrightarrow{\mathbf{MB'}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right|} = \frac{\left|\left|\overrightarrow{\mathbf{B''B'}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{BB'}}\right|\right|}, \qquad \frac{\left|\left|\overrightarrow{\mathbf{MB'}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right|} = \frac{\left|\left|\overrightarrow{\mathbf{MB''}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{AB}}\right|\right|} \implies \left|\left|\overrightarrow{\mathbf{MB''}}\right|\right| = (1-x)a,$$

$$\frac{\left|\left|\overrightarrow{\mathbf{BB''}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{BB'}}\right|\right|} = \frac{\left|\left|\overrightarrow{\mathbf{AM}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right|}, \qquad \frac{\left|\left|\overrightarrow{\mathbf{B''N}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{B'C'}}\right|\right|} = \frac{\left|\left|\overrightarrow{\mathbf{BB''}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{BB'}}\right|\right|} \implies \left|\left|\overrightarrow{\mathbf{B''N}}\right|\right| = xa.$$

Since $\left|\left|\overrightarrow{\mathbf{MN}}\right|\right| = \frac{\sqrt{5}}{3}a$, *by the Pythagorean Theorem,*

$$\left|\left|\overrightarrow{\mathbf{MN}}\right|\right|^2 = \left|\left|\overrightarrow{\mathbf{MB''}}\right|\right|^2 + \left|\left|\overrightarrow{\mathbf{B''N}}\right|\right|^2 \implies \frac{5}{9}a^2 = (1-x)^2a^2 + x^2a^2 \implies x \in \left\{\frac{1}{3}, \frac{2}{3}\right\}.$$

There are two possible positions for the segment, giving the solutions

$$\frac{\left|\left|\overrightarrow{\mathbf{AM}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right|} = \frac{\left|\left|\overrightarrow{\mathbf{BN}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{BC'}}\right|\right|} = \frac{1}{3}, \qquad \frac{\left|\left|\overrightarrow{\mathbf{AM}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{AB'}}\right|\right|} = \frac{\left|\left|\overrightarrow{\mathbf{BN}}\right|\right|}{\left|\left|\overrightarrow{\mathbf{BC'}}\right|\right|} = \frac{2}{3}.$$

◀

Homework

Problem 1.11.1 In a regular tetrahedron with vertices A, B, C, D and with $\left|\left|\overrightarrow{\mathbf{AB}}\right|\right| = a$, points M and N are the midpoints of the edges $[AB]$ and $[CD]$, respectively.

1. Find the length of the segment $[MN]$.
2. Find the angle between the lines $[MN]$ and $[BC]$.
3. Prove that $\overrightarrow{\mathbf{MN}} \perp \overrightarrow{\mathbf{AB}}$ and $\overrightarrow{\mathbf{MN}} \perp \overrightarrow{\mathbf{CD}}$.

Problem 1.11.2 In a tetrahedron $ABCD$, $\left|\left|\overrightarrow{\mathbf{AB}}\right|\right| = \left|\left|\overrightarrow{\mathbf{BC}}\right|\right|$, $\left|\left|\overrightarrow{\mathbf{AD}}\right|\right| = \left|\left|\overrightarrow{\mathbf{DC}}\right|\right|$. Prove that $\overrightarrow{\mathbf{AC}} \perp \overrightarrow{\mathbf{BD}}$.

Problem 1.11.3 In cube $\mathbf{ABCDD'C'B'A'}$ of edge of length a, find the distance between the lines that contain the diagonals $[A'B]$ and $[AC]$.

1.12 Cavalieri, and the Pappus-Guldin Rules

96 Theorem (Cavalieri's Principle) All planar regions with cross sections of proportional length at the same height have area in the same proportion. All solids with cross sections of proportional areas at the same height have their volume in the same proportion.

Proof: *We only provide the proof for the second statement, as the proof for the first is similar. Cut any two such solids by horizontal planes that produce cross sections of area* $A(x)$ *and* $cA(x)$, *where* $c > 0$ *is the constant of proportionality, at an arbitrary height* x *above a fixed base. From elementary calculus, we know that* $\int_{x_1}^{x_2} A(x)\mathrm{d}x$ *and* $\int_{x_1}^{x_2} cA(x)\mathrm{d}x$ *give the volume of the portion of each solid cut by all horizontal planes as* x *runs over some interval* $[x_1; x_2]$. *As* $\int_{x_1}^{x_2} A(x)\mathrm{d}x = c\int_{x_1}^{x_2} A(x)\mathrm{d}x$ *the corresponding volumes must also be proportional.* ❑

97 Example Use Cavalieri's Principle in order to deduce that the area enclosed by the ellipse with equation $\dfrac{x^2}{a^2}+\dfrac{y^2}{b}=1$, $a>0, b>0$, is πab.

Solution: ▶ *Consider the circle with equation $x^2+y^2=a^2$, as in figure 1.79. Then, for $y>0$,*

$$y=\sqrt{a^2-x^2}, \qquad y=\frac{b}{a}\sqrt{a^2-x^2}.$$

The corresponding ordinate for the ellipse and the circle are proportional, and hence, the corresponding chords for the ellipse and the circle will be proportional. By Cavalieri's first principle,

$$\begin{aligned}\text{Area of the ellipse} &= \frac{b}{a}\text{Area of the circle}\\ &= \frac{b}{a}\pi a^2\\ &= \pi ab.\end{aligned}$$

◀

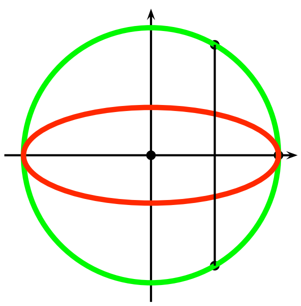

Figure 1.79: Ellipse and circle.

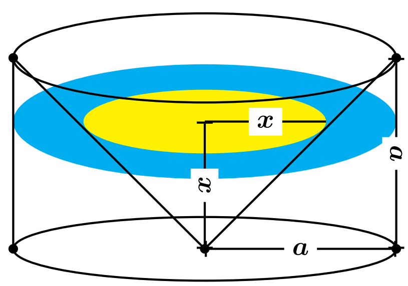

Figure 1.80: Punctured cylinder.

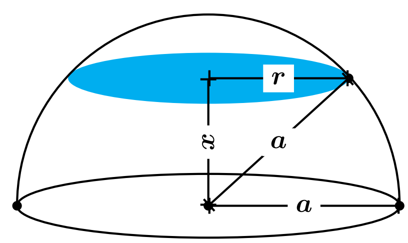

Figure 1.81: Hemisphere.

98 Example Use Cavalieri's Principle in order to deduce that the volume of a sphere with radius a is $\dfrac{4}{3}\pi a^3$.

Solution: ▶ *The following method is due to Archimedes, who was so proud of it that he wanted a sphere inscribed in a cylinder on his tombstone. We need to recall that the volume of a right circular cone with base radius a and height h is $\dfrac{\pi a^2 h}{3}$.*

Consider a hemisphere of radius a, as in figure 1.81. Cut a horizontal slice at height x, producing a circle of radius r. By the Pythagorean Theorem, $x^2+r^2=a^2$, and so this circular slab has area $\pi r^2=\pi(a^2-x^2)$. Now, consider a punctured cylinder of base radius a and height a, as in figure 1.80, with a cone of height a and base radius a cut from it. A horizontal slab at height x is an annular region of area $\pi a^2-\pi x^2$, which agrees with a horizontal slab for the sphere at the same height. By Cavalieri's Principle,

$$\begin{aligned}\text{Volume of the hemisphere} &= \text{Volume of the punctured cylinder}\\ &= \pi a^3-\frac{\pi a^3}{3}\\ &= \frac{2\pi a^3}{3}.\end{aligned}$$

It follows that the volume of the sphere is $2\left(\frac{2\pi a^3}{3}\right) = \frac{4\pi a^3}{3}$. ◀

Essentially the same method of proof as Cavalieri's Principle gives the next result.

99 Theorem (Pappus-Guldin Rule) The area of the lateral surface of a solid of revolution is equal to the product of the length of the generating curve on the side of the axis of revolution and the length of the path described by the centre of gravity of the generating curve under a full revolution. The volume of a solid of revolution is equal to the product of the area of the generating plane on one side of the revolution axis and the length of the path described by the centre of gravity of the area under a full revolution about the axis.

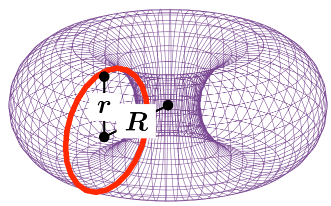

Figure 1.82: A torus.

100 Example Since the centre of gravity of a circle is at its centre, by the Pappus-Guldin Rule, the surface area of the torus with the generating circle having radius r, and radius of gyration R (as in figure 1.82) is $(2\pi r)(2\pi R) = 4\pi^2 rR$. Also, the volume of the solid torus is $(\pi r^2)(2\pi R) = 2\pi^2 r^2 R$.

Homework

Problem 1.12.1 Use the Pappus-Guldin Rule to find the lateral area and the volume of a right circular cone with base radius r and height h.

1.13 Dihedral Angles and Platonic Solids

101 Definition When two half planes intersect in space they intersect on a line. The portion of space bounded by the half planes and the line is called the *dihedral angle*. The intersecting line is called the *edge* of the dihedral angle and each of the two half planes of the dihedral angle is called a *face*. See figure 1.83.

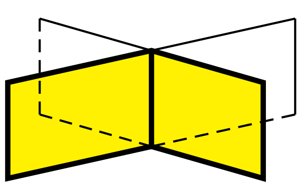

Figure 1.83: Dihedral Angles.

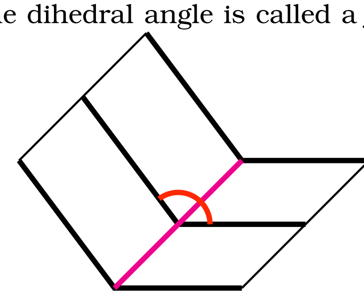

Figure 1.84: Rectilinear of a Dihedral Angle.

102 Definition The *rectilinear angle* of a dihedral angle is the angle whose sides are perpendicular to the edge of the dihedral angle at the same point, each on each of the faces. See figure 1.84.

All the rectilinear angles of a dihedral angle measure the same. Hence the *measure* of a dihedral angle is the measure of any one of its rectilinear angles.

In analogy to dihedral angles we now define polyhedral angles.

103 Definition The opening of three or more planes that meet at a common point is called a *polyhedral angle* or *solid angle*. In the particular case of three planes, we use the term *trihedral angle*. The common point is called the *vertex* of the polyhedral angle. Each of the intersecting lines of two consecutive planes is called an *edge* of the polyhedral angle. The portion of the planes lying between consecutive edges are called the *faces* of the polyhedral angle. The angles formed by adjacent edges are called *face angles*. A polyhedral angle is said to be *convex* if the section made by a plane cutting all its edges forms a convex polygon.

In the trihedral angle of figure 1.85, V is the vertex, $\triangle VAB$, $\triangle VBC$, $\triangle VCA$ are faces. Also, notice that in any polyhedral angle, any two adjacent faces form a dihedral angle.

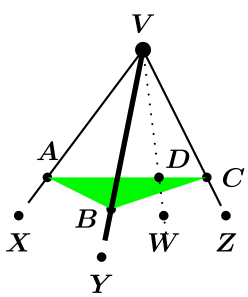

Figure 1.85: Trihedral Angle.

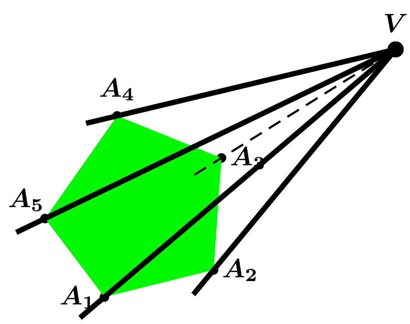

Figure 1.86: Polyhedral Angle.

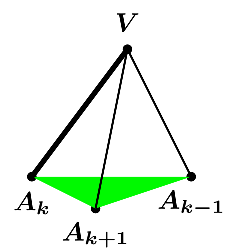

Figure 1.87: A, A_k, A_{k+1} are three consecutive vertices.

104 Theorem The sum of any two face angles of a trihedral angle is greater than the third face angle.

Proof: *Consider figure 1.85. If $\angle ZVX$ is smaller or equal to in size than either $\angle XVY$ or YVZ, then we are done, so assume that, say, $\angle ZVX > XVY$. We must demonstrate that*

$$\angle XVY + \angle YVZ > \angle ZVX.$$

Since we are assuming that $\angle ZVX > XVY$, we may draw, in $\angle XVY$ the line segment $[VW]$ such that $\angle XVW = \angle XVY$.

Through any point D of the segment $[VW]$, draw $\triangle ADC$ on the plane P containing the points V, X, Z. Take the point $B \in [VY]$ so that $VD = VB$. Consider now the plane containing the line segment $[AC]$ and the point B. Observe that $\triangle AVD \cong AVB$. Hence $AD = AB$. Now, by the triangle inequality in $\triangle ABC$, $AB + BC > CA$. This implies that $\angle BVC > \angle DVC$. Hence

$$\begin{aligned} \angle AVB + \angle BVC &= \angle AVD + \angle BVC \\ &> \angle AVD + \angle DVC \\ &= \angle AVC, \end{aligned}$$

which proves that $\angle XVY + \angle YVZ > \angle ZVX$, as wanted. ❑

105 Theorem The sum of the face angles of any convex polyhedral angle is less than 2π radians.

> **Proof:** *Let the polyhedral angle have n faces and vertex V. Let the faces be cut by a plane, intersecting the edges at the points $A_1, A_2, \ldots A_n$, say. An illustration can be seen in figure 1.86, where for convenience, we have depicted only five edges. Observe that the polygon $A_1A_2\cdots A_n$ is convex and that the sum of its interior angles is $\pi(n-2)$. We would like to prove that*
>
> $$\angle A_1VA_2 + \angle A_2VA_3 + \angle A_3VA_4 + \cdots + \angle A_{n-1}VA_n + \angle A_nVA_1 < 2\pi.$$
>
> *Now, let A_{k-1}, A_k, A_{k+1} be three consecutive vertices of of the polygon $A_1A_2\cdots A_n$. This notation means that $A_{k-1}A_kA_{k+1}$ represents any of the n triplets $A_1A_2A_3$, $A_2A_3A_4$, $A_3A_4A_5, \ldots$, $A_{n-2}A_{n-1}A_n$, $A_{n-1}A_nA_1$, $A_nA_1A_2$, that is, we let $A_0 = A_n$, $A_{n+1} = A_1$, $A_{n+2} = A_2$, etc. Consider the trihedral angle with vertex A_k and whose face angles at A_k are $\angle A_{k-1}A_kA_{k+1}$, $\angle VA_kA_{k-1}$, and $\angle VA_kA_{k+1}$, as in figure 1.87. Observe that as k ranges from 1 through n, the sum*
>
> $$\sum_{1\leq k\leq n} \angle A_{k-1}A_kA_{k+1} = \pi(n-2),$$
>
> *being the sum of the interior angles of the polygon $A_1A_2\cdots A_n$. By Theorem 104,*
>
> $$\angle VA_kA_{k-1} + \angle VA_kA_{k+1} > \angle A_{k-1}A_kA_{k+1}.$$
>
> *Thus*
>
> $$\sum_{1\leq k\leq n} VA_kA_{k-1} + \angle VA_kA_{k+1} > \sum_{1\leq k\leq n} \angle A_{k-1}A_kA_{k+1} = \pi(n-2).$$
>
> *Also,*
>
> $$\sum_{1\leq k\leq n} VA_kA_{k+1} + \angle VA_{k+1}A_k + \angle A_kVA_{k+1} = \pi n,$$
>
> *since this is summing the sum of the angles of the n triangles of the faces. But clearly*
>
> $$\sum_{1\leq k\leq n} VA_kA_{k+1} = \sum_{1\leq k\leq n} \angle VA_{k+1}A_k,$$
>
> *since one sum adds the angles in one direction and the other in the opposite direction. For the same reason,*
>
> $$\sum_{1\leq k\leq n} VA_kA_{k-1} = \sum_{1\leq k\leq n} \angle VA_kA_{k+1}.$$
>
> *Hence*
>
> $$\begin{aligned}\sum_{1\leq k\leq n} \angle A_kVA_{k+1} &= \pi n - \sum_{1\leq k\leq n} (VA_kA_{k+1} + \angle VA_{k+1}A_k) \\ &= \pi n - \sum_{1\leq k\leq n} (VA_kA_{k+1} + \angle VA_kA_{k-1}) \\ &< \pi n - \pi(n-2) \\ &= 2\pi,\end{aligned}$$
>
> *as we needed to shew.* ❑

106 Definition A *Platonic solid* is a polyhedron having congruent regular polygon as faces and having the same number of edges meeting at each corner.

Suppose a regular polygon with $n \geq 3$ sides is a face of a platonic solid with $m \geq 3$ faces meeting at a corner. Since each interior angle of this polygon measures $\dfrac{\pi(n-2)}{n}$, we must have in view of Theorem 105,

$$m\left(\frac{\pi(n-2)}{n}\right) < 2\pi \implies m(n-2) < 2n \implies (m-2)(n-2) < 4.$$

Since $n \geq 3$ and $m \geq 3$, the above inequality only holds for five pairs (n, m). Appealing to Euler's Formula for polyhedrons, which states that $V + F = E + 2$, where V is the number of vertices, F is the number of faces, and E is the number of edges of a polyhedron, we obtain the values in the following table.

m	n	S	E	F	Name of regular Polyhedron.
3	3	4	6	4	Tetrahedron or regular Pyramid.
4	3	8	12	6	Hexahedron or Cube.
3	4	6	12	8	Octahedron.
5	3	20	30	12	Dodecahedron.
3	5	12	30	20	Icosahedron.

Thus there are at most five Platonic solids. That there are exactly five can be seen by explicit construction. Figures 1.88 through 1.92 depict the Platonic solids.

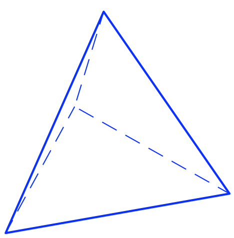

Figure 1.88: Tetrahedron.

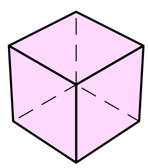

Figure 1.89: Cube or hexahedron.

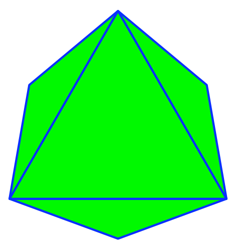

Figure 1.90: Octahedron.

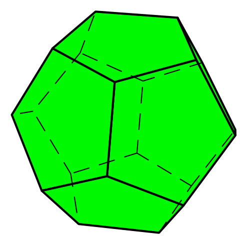

Figure 1.91: Dodecahedron

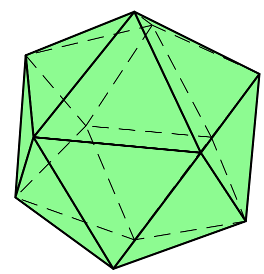

Figure 1.92: Icosahedron.

1.14 Spherical Trigonometry

Consider a point $B(x, y, z)$ in Cartesian coordinates. From $O(0, 0, 0)$ we draw a straight line to $B(x, y, z)$, and let its distance be ρ. We measure its inclination from the positive z-axis, let us say it is an angle of ϕ, $\phi \in [0; \pi]$ radians, as in figure 1.93. Observe that $z = \rho \cos \phi$. We now project the line segment $[OB]$ onto the xy-plane in order to find the polar coordinates of x and y. Let θ be angle that this projection makes with the positive x-axis. Since $OP = \rho \sin \phi$ we find $x = \rho \cos \theta \sin \phi$, $y = \rho \sin \theta \sin \phi$.

107 Definition Given a point (x, y, z) in Cartesian coordinates, its *spherical coordinates* are given by

$$x = \rho \cos \theta \sin \phi, \quad y = \rho \sin \theta \sin \phi, z = \rho \cos \phi.$$

Here ϕ is the *polar angle*, measured from the positive z-axis, and θ is the *azimuthal angle*, measured from the positive x-axis. By convention, $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi$.

Spherical coordinates are extremely useful when considering regions which are symmetric about a point.

108 Definition If a plane intersects with a sphere, the intersection will be a circle. If this circle contains the centre of the sphere, we call it a *great circle*. Otherwise we talk of a *small circle*. The *axis* of any circle on a sphere is the diameter of the sphere which is normal to the plane containing the circle. The endpoints of such a diameter are called the *poles* of the circle.

☞ *The radius of a great circle is the radius of the sphere. The poles of a great circle are equally distant from the plane of the circle, but this is not the case in a small circle. By* the pole *of a small circle, we mean the closest pole to the plane containing the circle. A pole of a circle is equidistant from every point of the circumference of the circle.*

109 Definition Given the centre of the sphere, and any two points of the surface of the sphere, a plane can be drawn. This plane will be unique if and only if the points are not diametrically opposite. In the case where the two points are not diametrically opposite, the great circle formed is split into a larger and a smaller arc by the two points. We call the smaller arc the *geodesic* joining the two points. If the two points are diametrically opposite then every plane containing the line forms with the sphere a great circle, and the arcs formed are then of equal length. In this case we take any such arc as a geodesic.

110 Definition A *spherical triangle* is a triangle on the surface of a sphere all whose vertices are connected by geodesics. The three arcs of great circles which form a spherical triangle are called the *sides* of the spherical triangle; the angles formed by the arcs at the points where they meet are called the *angles* of the spherical triangle.

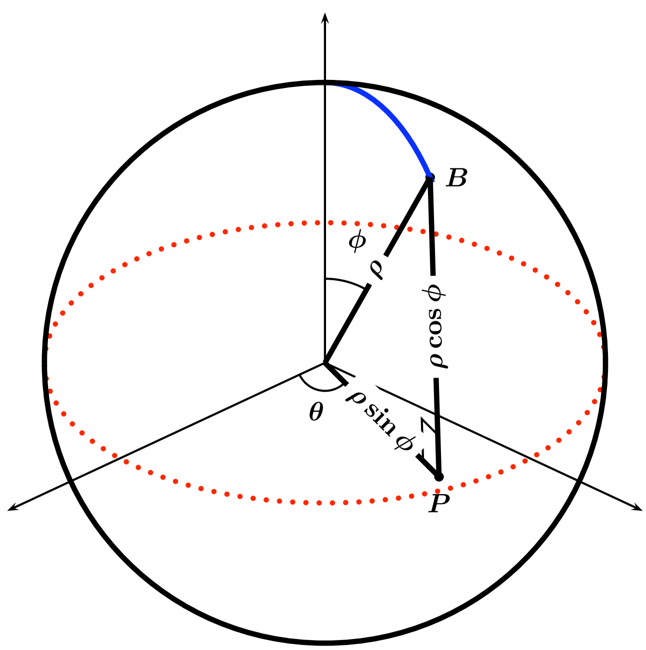

Figure 1.93: Spherical Coordinates.

If A, B, C are the vertices of a spherical triangle, it is customary to label the opposite arcs with the same letter name, but in lowercase.

☞ *A spherical triangle has then six angles: three vertex angles $\angle A$, $\angle B$, $\angle C$, and three arc angles, $\angle a$, $\angle b$, $\angle c$. Observe that if O is the centre of the sphere then*

$$\angle a = \angle\left(\overrightarrow{\mathbf{OB}}, \overrightarrow{\mathbf{OC}}\right), \quad \angle b = \angle\left(\overrightarrow{\mathbf{OC}}, \overrightarrow{\mathbf{OA}}\right), \quad \angle c = \angle\left(\overrightarrow{\mathbf{OA}}, \overrightarrow{\mathbf{OB}}\right),$$

and

$$\angle A = \angle\left(\overrightarrow{\mathbf{OA}} \times \overrightarrow{\mathbf{OB}}, \overrightarrow{\mathbf{OA}} \times \overrightarrow{\mathbf{OC}}\right), \quad \angle B = \angle\left(\overrightarrow{\mathbf{OB}} \times \overrightarrow{\mathbf{OC}}, \overrightarrow{\mathbf{OB}} \times \overrightarrow{\mathbf{OA}}\right), \quad \angle C = \angle\left(\overrightarrow{\mathbf{OC}} \times \overrightarrow{\mathbf{OA}}, \overrightarrow{\mathbf{OC}} \times \overrightarrow{\mathbf{OB}}\right).$$

111 Theorem Let $\triangle ABC$ be a spherical triangle. Then

$$\cos a \cos b + \sin a \sin b \cos C = \cos c.$$

Proof: *Consider a spherical triangle ABC with $A(x_1, y_1, z_1)$, $B(x_2, y_2, z_2)$, and let O be the centre and ρ be the radius of the sphere. In spherical coordinates this is, say,*

$$z_1 = \rho\cos\theta_1, \qquad x_1 = \rho\sin\theta_1\cos\phi_1, \qquad y_1 = \rho\sin\theta_1\sin\phi_1,$$
$$z_2 = \rho\cos\theta_2, \qquad x_2 = \rho\sin\theta_2\cos\phi_2, \qquad y_2 = \rho\sin\theta_2\sin\phi_2;$$

By a rotation we may assume that the z-axis passes through C. Then the following quantities give the square of the distance of the line segment $[AB]$:

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2, \quad \rho^2 + \rho^2 - 2\rho^2\cos\angle(AOB)\,.$$

Since $x_1^2 + y_1^2 + z_1^2 = \rho^2$, $x_2^2 + y_2^2 + z_2^2 = \rho^2$, we gather that

$$x_1x_2 + y_1y_2 + z_1z_2 = \rho^2\cos\angle(AOB)\,.$$

Therefore we obtain

$$\cos\theta_2\cos\theta_1 + \sin\theta_2\sin\theta_1\cos(\phi_1 - \phi_2) = \cos\angle(AOB)\,,$$

that is,

$$\cos a\cos b + \sin a\sin b\cos C = \cos c.$$

❑

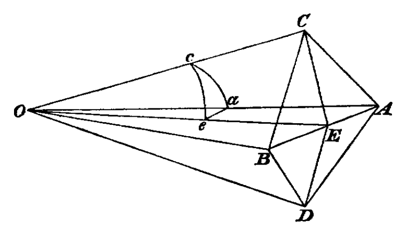

Figure 1.94: Theorem 112.

112 Theorem Let I be the dihedral angle of two adjacent faces of a regular polyhedron. Then

$$\sin\frac{I}{2} = \frac{\cos\frac{\pi}{n}}{\sin\frac{\pi}{m}}.$$

Proof: *See figure 1.94. Let AB be the edge common to the two adjacent faces, C and D the centres of the faces; bisect AB at E, and join CE and DE; CE and DE will be perpendicular to AB, and the angle CED is the angle of inclination of the two adjacent faces; we shall denote it by I. In the plane containing CE and DE draw CO and DO at right angles to CE and DE respectively, and meeting at O; about O as centre describe a sphere meeting OA, OC, OE at a, c, e respectively, so that cae forms a spherical triangle. Since AB is perpendicular to CE and DE, it is perpendicular to the plane CED, therefore the plane AOB which contains AB is perpendicular to the plane CED; hence the angle cea of the spherical triangle is a right angle. Let m be the number of sides in each face of the polyhedron, n the number of the plane angles which form each solid angle. Then the angle $ace = ACE = \dfrac{2\pi}{2m} = \dfrac{\pi}{m}$; and the angle cae is half one of the n equal angles formed on the sphere round a, that is, $cae = \dfrac{2\pi}{2n} = \dfrac{\pi}{n}$. From the right-angled triangle cae*

$$\cos cae = \cos cOe\sin ace,$$

that is
$$\cos\frac{\pi}{n} = \cos\left(\frac{\pi}{2} - \frac{I}{2}\right)\sin\frac{\pi}{m};$$

therefore
$$\sin\frac{I}{2} = \frac{\cos\frac{\pi}{n}}{\sin\frac{\pi}{m}}.$$

❑

113 Theorem Let r and R be, respectively, the radii of the inscribed and circumscribed spheres of a regular polyhedron. Then
$$r = \frac{a}{2}\cot\frac{\pi}{m}\tan\frac{I}{2}, \qquad R = \frac{a}{2}\tan\frac{I}{2}\tan\frac{\pi}{n}.$$
Here a is the length of any edge of the polyhedron, and I is the dihedral angle of any two faces.

Proof: *Let the edge $AB = a$, let $OC = r$ and $OA = R$, so that r is the radius of the inscribed sphere, and R is the radius of the circumscribed sphere. Then*
$$CE = AE\cot ACE = \frac{a}{2}\cot\frac{\pi}{m},$$
$$r = CE\tan CEO = CE\tan\frac{I}{2} = \frac{a}{2}\cot\frac{\pi}{m}\tan\frac{I}{2};$$
also
$$r = R\cos aOc = R\cot eca\cot eac = R\cot\frac{\pi}{m}\cot\frac{\pi}{n};$$
therefore
$$R = r\tan\frac{\pi}{m}\tan\frac{\pi}{n} = \frac{a}{2}\tan\frac{I}{2}\tan\frac{\pi}{n}.$$

❑

From the above formulæ we now easily find that the volume of the pyramid which has one face of the polyhedron for base and O for vertex is $\frac{r}{3}\cdot\frac{ma^2}{4}\cot\frac{\pi}{m}$, and therefore the volume of the polyhedron is $\frac{mFra^2}{12}\cot\frac{\pi}{m}$.

Furthermore, the area of one face of the polyhedron is $\frac{ma^2}{4}\cot\frac{\pi}{m}$, and therefore the surface area of the polyhedron is $\frac{mFa^2}{4}\cot\frac{\pi}{m}$.

Homework

Problem 1.14.1 The four vertices of a regular tetrahedron are
$$V_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad V_2 = \begin{pmatrix}-1/2\\ \sqrt{3}/2\\ 0\end{pmatrix},$$
$$V_3 = \begin{pmatrix}-1/2\\ -\sqrt{3}/2\\ 0\end{pmatrix}, \quad V_4 = \begin{pmatrix}0\\0\\ \sqrt{2}\end{pmatrix}.$$
What is the cosine of the dihedral angle between any pair of faces of the tetrahedron?

Problem 1.14.2 Consider a tetrahedron whose edge measures a. Shew that its volume is $\frac{a^3\sqrt{2}}{12}$, its surface area is $a^2\sqrt{3}$, and that the radius of the inscribed sphere is $\frac{a\sqrt{6}}{12}$.

Problem 1.14.3 Consider a cube whose edge measures a. Shew that its volume is a^3, its surface area is $6a^2$, and that the radius of the inscribed sphere is $\frac{a}{2}$.

Problem 1.14.4 Consider an octahedron whose edge

measures a. Shew that its volume is $\frac{a^3\sqrt{2}}{3}$, its surface area is $2a^2\sqrt{3}$, and that the radius of the inscribed sphere is $\frac{a\sqrt{6}}{6}$.

Problem 1.14.5 Consider a dodecahedron whose edge measures a. Shew that its volume is $\frac{a^3}{4}\left(15+7\sqrt{5}\right)$, its surface area is $3a^2\sqrt{25+10\sqrt{5}}$, and that the radius of the inscribed sphere is $\frac{a}{4}\sqrt{10+22\sqrt{\frac{1}{5}}}$.

Problem 1.14.6 Consider an icosahedron whose edge measures a. Shew that its volume is $\frac{5a^3}{12}\left(3+\sqrt{5}\right)$, its surface area is $5a^2\sqrt{3}$, and that the radius of the inscribed sphere is $\frac{a}{12}\left(5\sqrt{3}+\sqrt{15}\right)$.

1.15 Canonical Surfaces

In this section we consider various surfaces that we shall periodically encounter in subsequent sections. Just like in one-variable Calculus it is important to identify the equation and the shape of a line, a parabola, a circle, etc., it will become important for us to be able to identify certain families of often-occurring surfaces. We shall explore both their Cartesian and their parametric form. We remark that in order to parametrise curves ("one-dimensional entities") we needed one parameter, and that in order to parametrise surfaces we shall need to parameters.

Let us start with the plane. Recall that if a, b, c are real numbers, not all zero, then the Cartesian equation of a plane with normal vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ and passing through the point (x_0, y_0, z_0) is

$$a(x-x_0)+b(y-y_0)+c(z-z_0)=0.$$

If we know that the vectors $\vec{u}$ and $\vec{v}$ are on the plane (parallel to the plane) then with the parameters $p, q$the equation of the plane is

$$\begin{aligned} x-x_0 &= pu_1+qv_1, \\ y-y_0 &= pu_2+qv_2, \\ z-z_0 &= pu_3+qv_3. \end{aligned}$$

114 Definition A surface S consisting of all lines parallel to a given line Δ and passing through a given curve Γ is called a *cylinder*. The line Δ is called the *directrix* of the cylinder.

☞ *To recognise whether a given surface is a cylinder we look at its Cartesian equation. If it is of the form $f(A, B) = 0$, where A, B are secant planes, then the curve is a cylinder. Under these conditions, the lines generating S will be parallel to the line of equation $A = 0, B = 0$. In practice, if one of the variables x, y, or z is missing, then the surface is a cylinder, whose directrix will be the axis of the missing coordinate.*

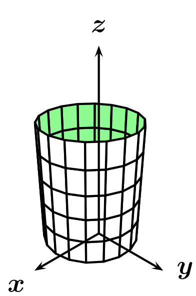

Figure 1.95: Circular cylinder $x^2 + y^2 = 1$.

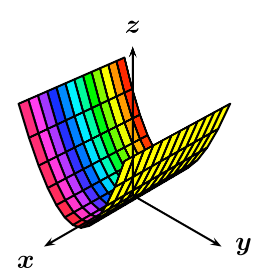

Figure 1.96: The parabolic cylinder $z = y^2$.

115 Example Figure 1.95 shews the cylinder with Cartesian equation $x^2 + y^2 = 1$. One starts with the circle $x^2 + y^2 = 1$ on the xy-plane and moves it up and down the z-axis. A parametrisation for this cylinder is the following:

$$x = \cos v, \qquad y = \sin v, \qquad z = u, \qquad u \in \mathbb{R}, v \in [0; 2\pi].$$

The Maple™ commands to graph this surface are:

```
> with(plots):
> implicitplot3d(x^2+y^2=1,x=-1..1,y=-1..1,z=-10..10);
> plot3d([cos(s),sin(s),t],s=-10..10,t=-10..10,numpoints=5001);
```

The method of parametrisation utilised above for the cylinder is quite useful when doing parametrisations in space. We refer to it as the method of *cylindrical coordinates*. In general, we first find the polar coordinates of x, y in the xy-plane, and then lift $(x, y, 0)$ parallel to the z-axis to (x, y, z):

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z.$$

See figure 1.97.

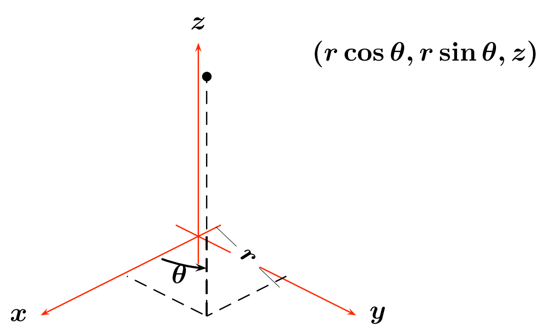

Figure 1.97: Cylindrical Coordinates.

116 Example Figure 1.96 shews the parabolic cylinder with Cartesian equation $z = y^2$. One starts with the parabola $z = y^2$ on the yz-plane and moves it up and down the x-axis. A parametrisation for this parabolic cylinder is the following:

$$x = u, \qquad y = v, \qquad z = v^2, \qquad u \in \mathbb{R}, v \in \mathbb{R}.$$

The Maple™ commands to graph this surface are:

```
> with(plots):
> implicitplot3d(z=y^2,x=-10..10,y=-10..10,z=-10..10,numpoints=5001);
> plot3d([t,s,s^2],s=-10..10,t=-10..10,numpoints=5001,axes=boxed);
```

117 Example Figure 1.98 shews the hyperbolic cylinder with Cartesian equation $x^2 - y^2 = 1$. One starts with the hyperbola $x^2 - y^2$ on the xy-plane and moves it up and down the z-axis. A parametrisation for this parabolic cylinder is the following:

$$x = \pm\cosh v, \qquad y = \sinh v, \qquad z = u, \qquad u \in \mathbb{R}, v \in \mathbb{R}.$$

We need a choice of sign for each of the portions. We have used the fact that $\cosh^2 v - \sinh^2 v = 1$. The Maple™ commands to graph this surface are:

```
> with(plots):
> implicitplot3d(x^2-y^2=1,x=-10..10,y=-10..10,z=-10..10,numpoints=5001);
> plot3d({[-cosh(s),sinh(s),t],[cosh(s),sinh(s),t]},
> s=-2..2,t=-10..10,numpoints=5001,axes=boxed);
```

118 Definition Given a point $\Omega \in \mathbb{R}^3$ (called the *apex*) and a curve Γ (called the generating curve), the surface S obtained by drawing rays from Ω and passing through Γ is called a *cone*.

In practice, if the Cartesian equation of a surface can be put into the form $f(\frac{A}{C}, \frac{B}{C}) = 0$, where A, B, C, are planes secant at exactly one point, then the surface is a cone, and its apex is given by $A = 0, B = 0, C = 0$.

119 Example The surface in $\mathbb{R}^3$ implicitly given by

$$z^2 = x^2 + y^2$$

is a cone, as its equation can be put in the form $\left(\frac{x}{z}\right)^2 + \left(\frac{y}{z}\right)^2 - 1 = 0$. Considering the planes $x = 0, y = 0, z = 0$, the apex is located at $(0, 0, 0)$. The graph is shewn in figure 1.100.

120 Definition A surface S obtained by making a curve Γ turn around a line Δ is called a *surface of revolution*. We then say that Δ is the axis of revolution. The intersection of S with a half-plane bounded by Δ is called a *meridian*.

If the Cartesian equation of S can be put in the form $f(A, \Sigma) = 0$, where A is a plane and Σ is a sphere, then the surface is of revolution. The axis of S is the line passing through the centre of Σ and perpendicular to the plane A.

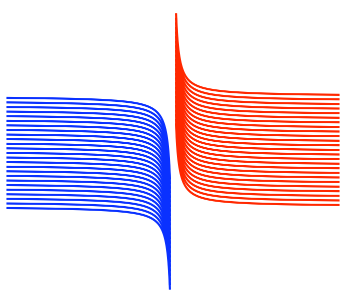

Figure 1.98: The hyperbolic cylinder $x^2 - y^2 = 1$.

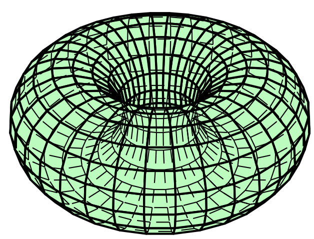

Figure 1.99: The torus.

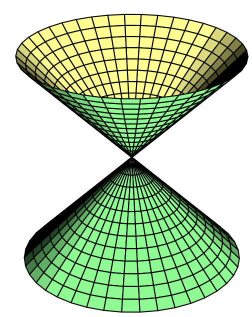

Figure 1.100: Cone $\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$.

121 Example Find the equation of the surface of revolution generated by revolving the hyperbola

$$x^2 - 4z^2 = 1$$

about the z-axis.

Solution: ► *Let (x, y, z) be a point on S. If this point were on the xz plane, it would be on the hyperbola, and its distance to the axis of rotation would be $|x| = \sqrt{1 + 4z^2}$. Anywhere else, the distance of (x, y, z) to the axis of rotation is the same as the distance of (x, y, z) to $(0, 0, z)$, that is $\sqrt{x^2 + y^2}$. We must have*

$$\sqrt{x^2 + y^2} = \sqrt{1 + 4z^2},$$

which is to say

$$x^2 + y^2 - 4z^2 = 1.$$

This surface is called a hyperboloid of one sheet. *See figure 1.104. Observe that when $z = 0$, $x^2 + y^2 = 1$ is a circle on the xy plane. When $x = 0$, $y^2 - 4z^2 = 1$ is a hyperbola on the yz plane. When $y = 0$, $x^2 - 4z^2 = 1$ is a hyperbola on the xz plane.*

A parametrisation for this hyperboloid is

$$x = \sqrt{1+4u^2}\cos v, \qquad y = \sqrt{1+4u^2}\sin v, \qquad z = u, \qquad u \in \mathbb{R}, v \in [0; 2\pi].$$

◀

122 Example The circle $(y-a)^2 + z^2 = r^2$, on the yz plane (a, r are positive real numbers) is revolved around the z-axis, forming a torus T. Find the equation of this torus.

Solution: ► *Let (x, y, z) be a point on T. If this point were on the yz plane, it would be on the circle, and the of the distance to the axis of rotation would be $y = a + \operatorname{sgn}(y-a)\sqrt{r^2 - z^2}$, where $\operatorname{sgn}(t)$ (with $\operatorname{sgn}(t) = -1$ if $t < 0$, $\operatorname{sgn}(t) = 1$ if $t > 0$, and $\operatorname{sgn}(0) = 0$) is the sign of t. Anywhere else, the distance from (x, y, z) to the z-axis is the distance of this point to the point (x, y, z) : $\sqrt{x^2 + y^2}$. We must have*

$$x^2 + y^2 = (a + \operatorname{sgn}(y-a)\sqrt{r^2 - z^2})^2 = a^2 + 2a\operatorname{sgn}(y-a)\sqrt{r^2 - z^2} + r^2 - z^2.$$

Rearranging

$$x^2 + y^2 + z^2 - a^2 - r^2 = 2a\operatorname{sgn}(y-a)\sqrt{r^2 - z^2},$$

or

$$(x^2 + y^2 + z^2 - (a^2 + r^2))^2 = 4a^2r^2 - 4a^2z^2$$

since $(\operatorname{sgn}(y-a))^2 = 1$, (it could not be 0, why?). Rearranging again,

$$(x^2 + y^2 + z^2)^2 - 2(a^2 + r^2)(x^2 + y^2) + 2(a^2 - r^2)z^2 + (a^2 - r^2)^2 = 0.$$

The equation of the torus thus, is of fourth degree, and its graph appears in figure 1.99.

A parametrisation for the torus generated by revolving the circle $(y-a)^2 + z^2 = r^2$ around the z-axis is

$$x = a\cos\theta + r\cos\theta\cos\alpha, \qquad y = a\sin\theta + r\sin\theta\cos\alpha, \qquad z = r\sin\alpha,$$

with $(\theta, \alpha) \in [-\pi; \pi]^2$.

◀

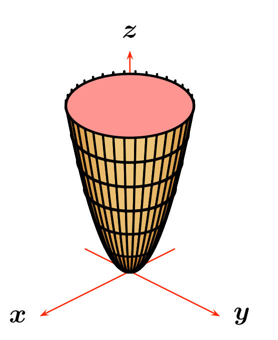

Figure 1.101: Paraboloid $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$.

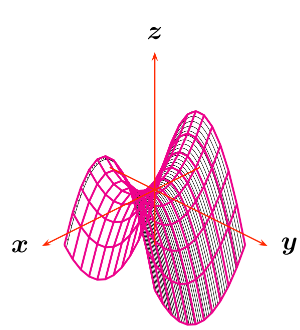

Figure 1.102: Hyperbolic paraboloid $z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$

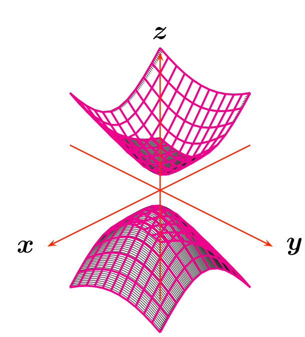

Figure 1.103: Two-sheet hyperboloid $\frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{b^2} + 1$.

123 Example The surface $z = x^2 + y^2$ is called an *elliptic paraboloid*. The equation clearly requires that $z \geq 0$. For fixed $z = c$, $c > 0$, $x^2 + y^2 = c$ is a circle. When $y = 0$, $z = x^2$ is a parabola on the xz plane. When $x = 0$, $z = y^2$ is a parabola on the yz plane. See figure 1.101. The following is a parametrisation of this paraboloid:

$$x = \sqrt{u}\cos v, \qquad y = \sqrt{u}\sin v, \qquad z = u, \qquad u \in [0; +\infty[, v \in [0; 2\pi].$$

124 Example The surface $z = x^2 - y^2$ is called a *hyperbolic paraboloid* or *saddle*. If $z = 0$, $x^2 - y^2 = 0$ is a pair of lines in the xy plane. When $y = 0$, $z = x^2$ is a parabola on the xz plane. When $x = 0$, $z = -y^2$ is a parabola on the yz plane. See figure 1.102. The following is a parametrisation of this hyperbolic paraboloid:

$$x = u, \qquad y = v, \qquad z = u^2 - v^2, \qquad u \in \mathbb{R}, v \in \mathbb{R}.$$

125 Example The surface $z^2 = x^2 + y^2 + 1$ is called an *hyperboloid of two sheets*. For $z^2 - 1 < 0$, $x^2 + y^2 < 0$ is impossible, and hence there is no graph when $-1 < z < 1$. When $y = 0$, $z^2 - x^2 = 1$ is a hyperbola on the xz plane. When $x = 0$, $z^2 - y^2 = 1$ is a hyperbola on the yz plane. When $z = c$ is a constant $c > 1$, then the $x^2 + y^2 = c^2 - 1$ are circles. See figure 1.103. The following is a parametrisation for the top sheet of this hyperboloid of two sheets

$$x = u\cos v, \qquad y = u\sin v, \qquad z = u^2 + 1, \qquad u \in \mathbb{R}, v \in [0; 2\pi]$$

and the following parametrises the bottom sheet,

$$x = u\cos v, \qquad y = u\sin v, \qquad z = -u^2 - 1, \qquad u \in \mathbb{R}, v \in [0; 2\pi],$$

126 Example The surface $z^2 = x^2 + y^2 - 1$ is called an *hyperboloid of one sheet*. For $x^2 + y^2 < 1$, $z^2 < 0$ is impossible, and hence there is no graph when $x^2 + y^2 < 1$. When $y = 0$, $z^2 - x^2 = -1$ is a hyperbola on the xz plane. When $x = 0$, $z^2 - y^2 = -1$ is a hyperbola on the yz plane. When $z = c$ is a constant, then the $x^2 + y^2 = c^2 + 1$ are circles See figure 1.104. The following is a parametrisation for this hyperboloid of one sheet

$$x = \sqrt{u^2+1}\cos v, \qquad y = \sqrt{u^2+1}\sin v, \qquad z = u, \qquad u \in \mathbb{R}, v \in [0; 2\pi],$$

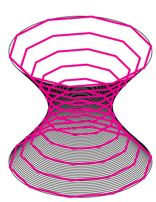

Figure 1.104: One-sheet hyperboloid $\dfrac{z^2}{c^2} = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - 1$.

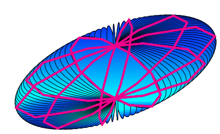

Figure 1.105: Ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$.

127 Example Let a, b, c be strictly positive real numbers. The surface $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ is called an *ellipsoid*. For $z = 0$, $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}1$ is an ellipse on the xy plane.When $y = 0$, $\dfrac{x^2}{a^2} + \dfrac{z^2}{c^2} = 1$ is an ellipse on the xz plane. When $x = 0$, $\dfrac{z^2}{c^2} + \dfrac{y^2}{b^2} = 1$ is an ellipse on the yz plane. See figure 1.105. We may parametrise the ellipsoid using spherical coordinates:

$$x = a\cos\theta\sin\phi, \qquad y = b\sin\theta\sin\phi, \qquad z = c\cos\phi, \qquad \theta \in [0; 2\pi], \phi \in [0; \pi].$$

Homework

Problem 1.15.1 Find the equation of the surface of revolution S generated by revolving the ellipse $4x^2 + z^2 = 1$ about the z-axis.

Problem 1.15.2 Find the equation of the surface of revolution generated by revolving the line $3x + 4y = 1$ about the y-axis .

Problem 1.15.3 Describe the surface parametrised by $\varphi(u, v) \mapsto (v \cos u, v \sin u, au)$, $(u, v) \in (0, 2\pi) \times (0, 1)$, $a > 0$.

Problem 1.15.4 Describe the surface parametrised by $\varphi(u, v) = (au \cos v, bu \sin v, u^2)$, $(u, v) \in (1, +\infty) \times (0, 2\pi)$, $a, b > 0$.

Problem 1.15.5 Consider the spherical cap defined by

$$S = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1,\ z \geq 1/\sqrt{2}\}.$$

Parametrise S using Cartesian, Spherical, and Cylindrical coordinates.

Problem 1.15.6 Demonstrate that the surface in $\mathbb{R}^3$

$$S : e^{x^2+y^2+z^2} - (x + z)e^{-2xz} = 0,$$

implicitly defined, is a cylinder.

Problem 1.15.7 Shew that the surface in $\mathbb{R}^3$ implicitly defined by

$$x^4 + y^4 + z^4 - 4xyz(x + y + z) = 1$$

is a surface of revolution, and find its axis of revolution.

Problem 1.15.8 Shew that the surface S in $\mathbb{R}^3$ given implicitly by the equation

$$\frac{1}{x-y} + \frac{1}{y-z} + \frac{1}{z-x} = 1$$

is a cylinder and find the direction of its directrix.

Problem 1.15.9 Shew that the surface S in $\mathbb{R}^3$ implicitly defined as

$$xy + yz + zx + x + y + z + 1 = 0$$

is of revolution and find its axis.

Problem 1.15.10 Demonstrate that the surface in $\mathbb{R}^3$ given implicitly by

$$z^2 - xy = 2z - 1$$

is a cone

Problem 1.15.11 (Putnam Exam 1970) Determine, with proof, the radius of the largest circle which can lie on the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \qquad a > b > c > 0.$$

Problem 1.15.12 The hyperboloid of one sheet in figure 1.106 has the property that if it is cut by planes at $z = \pm 2$, its projection on the xy plane produces the ellipse $x^2 + \frac{y^2}{4} = 1$, and if it is cut by a plane at $z = 0$, its projection on the xy plane produces the ellipse $4x^2 + y^2 = 1$. Find its equation.

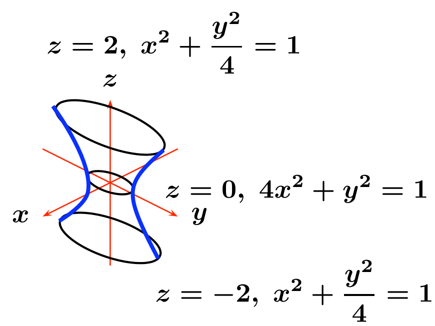

Figure 1.106: Problem 1.15.12.

1.16 Parametric Curves in Space

In analogy to curves on the plane, we now define curves in space.

128 Definition Let $[a; b] \subseteq \mathbb{R}$. A *parametric curve* representation r of a curve Γ is a function $\mathrm{r} : [a; b] \to \mathbb{R}^3$, with

$$\mathrm{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix},$$

and such that $\mathbf{r}([a;b]) = \Gamma$. $\mathbf{r}(a)$ is the *initial point* of the curve and $\mathbf{r}(b)$ its *terminal point*. A curve is *closed* if its initial point and its final point coincide. The *trace* of the curve $\mathbf{r}$ is the set of all images of $\mathbf{r}$, that is, Γ. The length of the curve is

$$\int_{\Gamma} \left|\left|\mathrm{d}\overrightarrow{\mathrm{r}}\right|\right|.$$

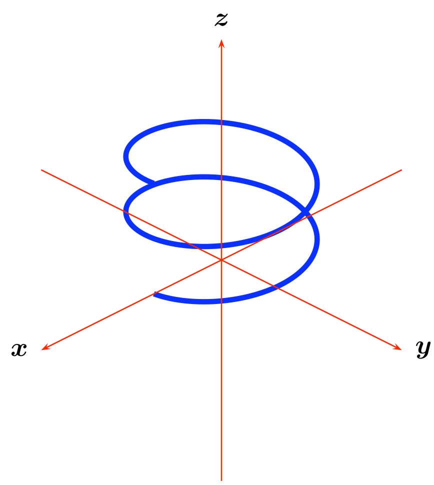

Figure 1.107: Helix.

129 Example The trace of

$$\mathrm{r}(t) = \overrightarrow{\mathrm{i}}\cos t + \overrightarrow{\mathrm{j}}\sin t + \overrightarrow{\mathrm{k}}\,t$$

is known as a *cylindrical helix.* To find the length of the helix as t traverses the interval $[0; 2\pi]$, first observe that

$$\left|\left|\mathrm{d}\overrightarrow{\mathrm{x}}\right|\right| = \left|\left|(\sin t)^2 + (-\cos t)^2 + 1\right|\right|\mathrm{d}t = \sqrt{2}\mathrm{d}t,$$

and thus the length is

$$\int_0^{2\pi} \sqrt{2}\mathrm{d}t = 2\pi\sqrt{2}.$$

The Maple™ commands to graph this curve and to find its length are:

```
>   with(plots):
>   with(Student[VectorCalculus]):
>   spacecurve([cos(t),sin(t),t],t=0..2*Pi,axes=normal);
>   PathInt(1,[x,y,z]=Path(<cos(t),sin(t),t>,0..2*Pi));
```

130 Example Find a parametric representation for the curve resulting by the intersection of the plane $3x + y + z = 1$ and the cylinder $x^2 + 2y^2 = 1$ in $\mathbb{R}^3$.

Solution: ► *The projection of the intersection of the plane $3x + y + z = 1$ and the cylinder is the ellipse $x^2 + 2y^2 = 1$, on the xy-plane. This ellipse can be parametrised as*

$$x = \cos t,\ y = \frac{\sqrt{2}}{2}\sin t, \qquad 0 \leq t \leq 2\pi.$$

From the equation of the plane,

$$z = 1 - 3x - y = 1 - 3\cos t - \frac{\sqrt{2}}{2}\sin t.$$

Thus we may take the parametrisation

$$\mathbf{r}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} \cos t \\ \dfrac{\sqrt{2}}{2}\sin t \\ 1 - 3\cos t - \dfrac{\sqrt{2}}{2}\sin t \end{bmatrix}.$$

◀

131 Example Let a, b, c be strictly positive real numbers. Consider the the region

$$\mathscr{R} = \left\{(x, y, z) \in \mathbb{R}^3 : |x| \leq a, |y| \leq b, z = c\right\}.$$

A point P moves along the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \qquad z = c + 1$$

once around, and acts as a source light projecting a shadow of $\mathscr{R}$ onto the xy-plane. Find the area of this shadow.

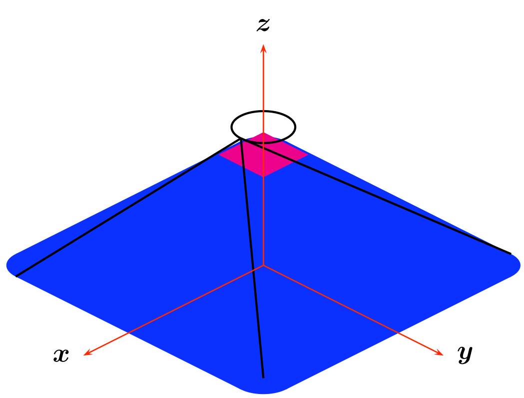

Figure 1.108: Problem 1.16.4.

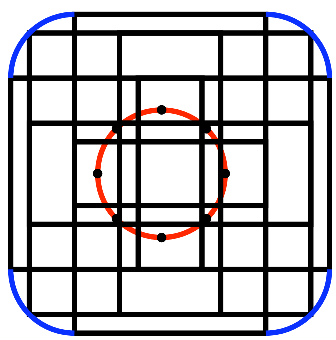
Figure 1.109: Problem 1.16.4.

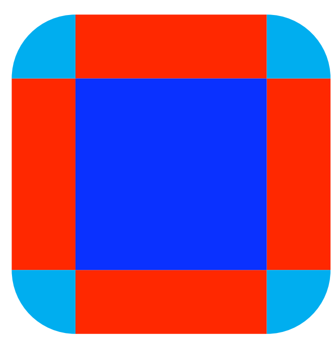
Figure 1.110: Problem 1.16.4.

Solution: ▶ *First consider the same problem as P moves around the circle*

$$x^2 + y^2 = 1, \quad z = c + 1$$

and the region is $\mathcal{R}' = \left\{(x, y, z) \in \mathbb{R}^3 : |x| \leq 1, |y| \leq 1, z = c\right\}$.

For fixed P $(u, v, c + 1)$ on the circle, the image of $\mathcal{R}'$ (a 2×2 square) on the xy plane is a $(2c + 2) \times (2c + 2)$ square with centre at the point $Q(-cu, -cv, 0)$ (figure 1.109). As P moves along the circle, Q moves along the circle with equation $x^2 + y^2 = c^2$ on the xy-plane (figure 1.109), being the centre of a $(2c+2) \times (2c+2)$ square. This creates a region as in figure 1.110, where each quarter circle has radius c, and the central square has side $2c + 2$, of area

$$\pi c^2 + 4(c + 1)^2 + 8c(c + 1).$$

Resizing to a region

$$\mathcal{R} = \left\{(x, y, z) \in \mathbb{R}^3 : |x| \leq a, |y| \leq b, z = c\right\},$$

and an ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad z = c + 1$$

we use instead of $c + 1$, $a(c + 1)$ *(parallel to the* x*-axis) and* $b(c + 1)$ *(parallel to the* y*-axis), so that the area shadowed is*

$$\pi ab(c + 1)^2 + 4ab(c + 1)^2 + 4abc(c + 1) = c^2 ab(\pi + 12) + 16abc + 4ab.$$

◀

Homework

Problem 1.16.1 Let $\mathscr{C}$ be the curve in $\mathbb{R}^3$ defined by

$$x = t^2, \qquad y = 4t^{3/2}, \qquad z = 9t, \qquad t \in [0; +\infty[.$$

Calculate the distance along $\mathscr{C}$ from $(1, 4, 9)$ to $(16, 32, 36)$.

Problem 1.16.2 Consider the surfaces in $\mathbb{R}^3$ implicitly defined by

$$z - x^2 - y^2 - 1 = 0, \qquad z + x^2 + y^2 - 3 = 0.$$

Describe, as vividly as possible these surfaces and their intersection, if they at all intersect. Find a parametric equation for the curve on which they intersect, if they at all intersect.

Problem 1.16.3 Consider the space curve

$$\overrightarrow{r} : t \mapsto \begin{bmatrix} \dfrac{t^4}{1 + t^2} \\ \dfrac{t^3}{1 + t^2} \\ \dfrac{t^2}{1 + t^2} \end{bmatrix}.$$

Let $t_k, 1 \leq k \leq 4$ be non-zero real numbers. Prove that $\overrightarrow{r}(t_1)$, $\overrightarrow{r}(t_2)$, $\overrightarrow{r}(t_3)$, and $\overrightarrow{r}(t_4)$ are coplanar if and only if

$$\frac{1}{t_1} + \frac{1}{t_2} + \frac{1}{t_3} + \frac{1}{t_4} = 0.$$

Problem 1.16.4 Give a parametrisation for the part of the ellipsoid

$$x^2 + \frac{y^2}{9} + \frac{z^2}{4} = 1$$

which lies on top of the plane $x + y + z = 0$.

Problem 1.16.5 Let P be the point $(2, 0, 1)$ and consider the curve $\mathscr{C} : z = y^2$ on the yz-plane. As a point Q moves along $\mathscr{C}$, let R be the point of intersection of $\overleftrightarrow{PQ}$ and the xy-plane. Graph all points R on the xy-plane.

Problem 1.16.6 Let a be a real number parameter, and consider the planes

$$P_1 : ax + y + z = -a,$$

$$P_2 : x - ay + az = -1.$$

Let l be their intersection line.

1. Find a direction vector for l.
2. As a varies through $\mathbb{R}$, l describes a surface $\mathcal{S}$ in $\mathbb{R}^3$. Let (x, y, z) be the point of intersection of this surface and the plane $z = c$. Find an equation relating x and y.
3. Find the volume bounded by the two planes, $x = 0$, and $x = 1$, and the surface $\mathcal{S}$ as c varies.

1.17 Multidimensional Vectors

We briefly describe space in n-dimensions. The ideas expounded earlier about the plane and space carry almost without change.

132 Definition $\mathbb{R}^n$ is the n-dimensional space, the collection

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_k \in \mathbb{R} \right\}.$$

133 Definition If $\overrightarrow{a}$ and $\overrightarrow{b}$ are two vectors in $\mathbb{R}^n$ their *vector sum* $\overrightarrow{a}+\overrightarrow{b}$ is defined by the coordinatewise addition

$$\overrightarrow{a}+\overrightarrow{b} = \begin{bmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{bmatrix}. \tag{1.20}$$

134 Definition A real number $\alpha \in \mathbb{R}$ will be called a *scalar*. If $\alpha \in \mathbb{R}$ and $\overrightarrow{a} \in \mathbb{R}^n$ we define *scalar multiplication* of a vector and a scalar by the coordinatewise multiplication

$$\alpha\overrightarrow{a} = \begin{bmatrix} \alpha a_1 \\ \alpha a_2 \\ \vdots \\ \alpha a_n \end{bmatrix}. \tag{1.21}$$

135 Definition The *standard ordered basis* for $\mathbb{R}^n$ is the collection of vectors

$$\{\overrightarrow{e}_1, \overrightarrow{e}_2, \ldots, \overrightarrow{e}_n, \}$$

with

$$\overrightarrow{e}_k = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

(a 1 in the k slot and 0's everywhere else). Observe that

$$\sum_{k=1}^{n} \alpha_k \overrightarrow{e}_k = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

136 Definition Given vectors $\overrightarrow{a}$, $\overrightarrow{b}$ of $\mathbb{R}^n$, their *dot product* is

$$\overrightarrow{a}\bullet\overrightarrow{b} = \sum_{k=1}^{n} a_k b_k.$$

We now establish one of the most useful inequalities in analysis.

137 Theorem (Cauchy-Bunyakovsky-Schwarz Inequality) Let $\overrightarrow{x}$ and $\overrightarrow{y}$ be any two vectors in $\mathbb{R}^n$. Then we have

$$|\overrightarrow{x}\bullet\overrightarrow{y}| \leq ||\overrightarrow{x}||||\overrightarrow{y}||.$$

Proof: *Since the norm of any vector is non-negative, we have*

$$\begin{aligned} ||\overrightarrow{x}+t\overrightarrow{y}|| \geq 0 \quad &\iff \quad (\overrightarrow{x}+t\overrightarrow{y})\bullet(\overrightarrow{x}+t\overrightarrow{y}) \geq 0 \\ &\iff \quad \overrightarrow{x}\bullet\overrightarrow{x} + 2t\overrightarrow{x}\bullet\overrightarrow{y} + t^2\overrightarrow{y}\bullet\overrightarrow{y} \geq 0 \\ &\iff \quad ||\overrightarrow{x}||^2 + 2t\overrightarrow{x}\bullet\overrightarrow{y} + t^2||\overrightarrow{y}||^2 \geq 0. \end{aligned}$$

This last expression is a quadratic polynomial in t which is always non-negative. As such its discriminant must be non-positive, that is,

$$(2\overrightarrow{x}\bullet\overrightarrow{y})^2 - 4(\left|\left|\overrightarrow{x}\right|\right|^2)(\left|\left|\overrightarrow{y}\right|\right|^2) \leq 0 \iff |\overrightarrow{x}\bullet\overrightarrow{y}| \leq \left|\left|\overrightarrow{x}\right|\right|\left|\left|\overrightarrow{y}\right|\right|,$$

giving the theorem. ❑

☞ *The above proof works not just for* $\mathbb{R}^n$ *but for any* vector space *(cf. below) that has an inner product.*

The form of the Cauchy-Bunyakovsky-Schwarz most useful to us will be

$$\left|\sum_{k=1}^{n} x_k y_k\right| \leq \left(\sum_{k=1}^{n} x_k^2\right)^{1/2} \left(\sum_{k=1}^{n} y_k^2\right)^{1/2}, \tag{1.22}$$

for real numbers x_k, y_k.

138 Corollary (Triangle Inequality) Let $\overrightarrow{a}$ and $\overrightarrow{b}$ be any two vectors in $\mathbb{R}^n$. Then we have

$$\left|\left|\overrightarrow{a} + \overrightarrow{b}\right|\right| \leq \left|\left|\overrightarrow{a}\right|\right| + \left|\left|\overrightarrow{b}\right|\right|.$$

Proof:

$$\begin{aligned} \left|\left|\overrightarrow{a} + \overrightarrow{b}\right|\right|^2 &= (\overrightarrow{a} + \overrightarrow{b})\bullet(\overrightarrow{a} + \overrightarrow{b}) \\ &= \overrightarrow{a}\bullet\overrightarrow{a} + 2\overrightarrow{a}\bullet\overrightarrow{b} + \overrightarrow{b}\bullet\overrightarrow{b} \\ &\leq \left|\left|\overrightarrow{a}\right|\right|^2 + 2\left|\left|\overrightarrow{a}\right|\right|\left|\left|\overrightarrow{b}\right|\right| + \left|\left|\overrightarrow{b}\right|\right|^2 \\ &= (\left|\left|\overrightarrow{a}\right|\right| + \left|\left|\overrightarrow{b}\right|\right|)^2, \end{aligned}$$

from where the desired result follows. ❑

Again, the preceding proof is valid in any vector space that has a norm.

139 Definition Let $\overrightarrow{x}$ and $\overrightarrow{y}$ be two non-zero vectors in a vector space over the real numbers. Then the angle $(\widehat{\overrightarrow{x}, \overrightarrow{y}})$ between them is given by the relation

$$\cos(\widehat{\overrightarrow{x}, \overrightarrow{y}}) = \frac{\overrightarrow{x}\bullet\overrightarrow{y}}{\left|\left|\overrightarrow{x}\right|\right|\left|\left|\overrightarrow{y}\right|\right|}.$$

This expression agrees with the geometry in the case of the dot product for $\mathbb{R}^2$ and $\mathbb{R}^3$.

140 Example Assume that $a_k, b_k, c_k, k = 1, \ldots, n$, are positive real numbers. Shew that

$$\left(\sum_{k=1}^{n} a_k b_k c_k\right)^4 \leq \left(\sum_{k=1}^{n} a_k^4\right)\left(\sum_{k=1}^{n} b_k^4\right)\left(\sum_{k=1}^{n} c_k^2\right)^2.$$

Solution: ▶ *Using CBS on* $\sum_{k=1}^{n}(a_k b_k)c_k$ *once we obtain*

$$\sum_{k=1}^{n} a_k b_k c_k \leq \left(\sum_{k=1}^{n} a_k^2 b_k^2\right)^{1/2} \left(\sum_{k=1}^{n} c_k^2\right)^{1/2}.$$

Using CBS again on $\left(\sum_{k=1}^{n} a_k^2 b_k^2\right)^{1/2}$ *we obtain*

$$\begin{aligned} \sum_{k=1}^{n} a_k b_k c_k &\leq \left(\sum_{k=1}^{n} a_k^2 b_k^2\right)^{1/2} \left(\sum_{k=1}^{n} c_k^2\right)^{1/2} \\ &\leq \left(\sum_{k=1}^{n} a_k^4\right)^{1/4} \left(\sum_{k=1}^{n} b_k^4\right)^{1/4} \left(\sum_{k=1}^{n} c_k^2\right)^{1/2}, \end{aligned}$$

which gives the required inequality. ◀

We now use the CBS inequality to establish another important inequality. We need some preparatory work.

141 Lemma Let $a_k > 0$, $q_k > 0$, with $\sum_{k=1}^{n} q_k = 1$. Then

$$\lim_{x\to 0} \log\left(\sum_{k=1}^{n} q_k a_k^x\right)^{1/x} = \sum_{k=1}^{n} q_k \log a_k.$$

Proof: *Recall that* $\log(1+x) \sim x$ *as* $x \to 0$. *Thus*

$$\begin{aligned} \lim_{x\to 0} \log\left(\sum_{k=1}^{n} q_k a_k^x\right)^{1/x} &= \lim_{x\to 0} \frac{\log\left(\sum_{k=1}^{n} q_k a_k^x\right)}{x} \\ &= \lim_{x\to 0} \frac{\sum_{k=1}^{n} q_k (a_k^x - 1)}{x} \\ &= \lim_{x\to 0} \sum_{k=1}^{n} q_k \frac{(a_k^x - 1)}{x} \\ &= \sum_{k=1}^{n} q_k \log a_k. \end{aligned}$$

❑

142 Theorem (Arithmetic Mean-Geometric Mean Inequality) Let $a_k \geq 0$. Then

$$\sqrt[n]{a_1 a_2 \cdots a_n} \leq \frac{a_1 + a_2 + \cdot + a_n}{n}.$$

Proof: *If* $b_k \geq 0$, *then by CBS*

$$\frac{1}{n}\sum_{k=1}^{n} b_k \geq \left(\frac{1}{n}\sum_{k=1}^{n} \sqrt{b_k}\right)^2. \tag{1.23}$$

Successive applications of (1.23) yield the monotone decreasing sequence

$$\frac{1}{n}\sum_{k=1}^{n} a_k \geq \left(\frac{1}{n}\sum_{k=1}^{n} \sqrt{a_k}\right)^2 \geq \left(\frac{1}{n}\sum_{k=1}^{n} \sqrt[4]{a_k}\right)^4 \geq \cdots,$$

which by Lemma 141 has limit

$$\exp\left(\frac{1}{n}\sum_{k=1}^{n} \log a_k\right) = \sqrt[n]{a_1 a_2 \cdots a_n},$$

giving

$$\sqrt[n]{a_1 a_2 \cdots a_n} \leq \frac{a_1 + a_2 + \cdot + a_n}{n},$$

as wanted. ❑

143 Example For any positive integer $n > 1$ we have

$$1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2n-1) < n^n.$$

For, by AMGM,

$$1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2n-1) < \left(\frac{1+3+5+\cdots+(2n-1)}{n}\right)^n = \left(\frac{n^2}{n}\right)^n = n^n.$$

Notice that since the factors are unequal we have strict inequality.

144 Definition Let $a_1 > 0, a_2 > 0, \ldots, a_n > 0$. Their *harmonic mean* is given by

$$\frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}}.$$

As a corollary to AMGM we obtain

145 Corollary (Harmonic Mean-Geometric Mean Inequality) Let $b_1 > 0, b_2 > 0, \ldots, b_n > 0$. Then

$$\frac{n}{\frac{1}{b_1} + \frac{1}{b_2} + \cdots + \frac{1}{b_n}} \leq (b_1 b_2 \cdots b_n)^{1/n}.$$

Proof: *This follows by putting* $a_k = \dfrac{1}{b_k}$ *in Theorem 142 . For then*

$$\left(\frac{1}{b_1}\frac{1}{b_2}\cdots\frac{1}{b_n}\right)^{1/n} \leq \frac{\frac{1}{b_1} + \frac{1}{b_2} + \cdots + \frac{1}{b_n}}{n}.$$

❑

Combining Theorem 142 and Corollary 145, we deduce

146 Corollary (Harmonic Mean-Arithmetic Mean Inequality) Let $b_1 > 0, b_2 > 0, \ldots, b_n > 0$. Then

$$\frac{n}{\frac{1}{b_1} + \frac{1}{b_2} + \cdots + \frac{1}{b_n}} \leq \frac{b_1 + b_2 + \cdots + b_n}{n}.$$

147 Example Let $a_k > 0$, and $s = a_1 + a_2 + \cdots + a_n$. Prove that

$$\sum_{k=1}^{n} \frac{s}{s - a_k} \geq \frac{n^2}{n-1}$$

and

$$\sum_{k=1}^{n} \frac{a_k}{s - a_k} \geq \frac{n}{n-1}.$$

Solution: ► *Put* $b_k = \dfrac{s}{s - a_k}$. *Then*

$$\sum_{k=1}^{n} \frac{1}{b_k} = \sum_{k=1}^{n} \frac{s - a_k}{s} = n - 1$$

and from Corollary 146,

$$\frac{n}{n-1} \leq \frac{\sum_{k=1}^{n} \frac{s}{s - a_k}}{n},$$

from where the first inequality is proved.

Since $\dfrac{s}{s-a_k} - 1 = \dfrac{a_k}{s-a_k}$, *we have*

$$\begin{aligned}\sum_{k=1}^{n} \frac{a_k}{s-a_k} &= \sum_{k=1}^{n} \left(\frac{s}{s-a_k} - 1\right) \\ &= \sum_{k=1}^{n} \left(\frac{s}{s-a_k}\right) - n \\ &\geq \frac{n^2}{n-1} - n \\ &= \frac{n}{n-1}.\end{aligned}$$

◀

Homework

Problem 1.17.1 The *Arithmetic Mean Geometric Mean Inequality* says that if $a_k \geq 0$ then

$$(a_1 a_2 \cdots a_n)^{1/n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

Equality occurs if and only if $a_1 = a_2 = \ldots = a_n$. In this exercise you will follow the steps of a proof by George Pólya.

1. Prove that $\forall x \in \mathbb{R},\ x \leq e^{x-1}$.
2. Put
$$A_k = \frac{n a_k}{a_1 + a_2 + \cdots + a_n},$$
and $G_n = a_1 a_2 \cdots a_n$. Prove that
$$A_1 A_2 \cdots A_n = \frac{n^n G_n}{(a_1 + a_2 + \cdots + a_n)^n},$$
and that
$$A_1 + A_2 + \cdots + A_n = n.$$
3. Deduce that
$$G_n \leq \left(\frac{a_1 + a_2 + \cdots + a_n}{n}\right)^n.$$
4. Prove the AMGM inequality by assembling the results above.

Problem 1.17.2 Demonstrate that if $x_1, x_2, \ldots, x_n$, are strictly positive real numbers then

$$(x_1 + x_2 + \ldots + x_n)\left(\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}\right) \geq n^2.$$

Problem 1.17.3 (USAMO 1978) Let a, b, c, d, e be real numbers such that

$$a + b + c + d + e = 8, \qquad a^2 + b^2 + c^2 + d^2 + e^2 = 16.$$

Maximise the value of e.

Problem 1.17.4 Find all positive real numbers

$$a_1 \leq a_2 \leq \ldots \leq a_n$$

such that

$$\sum_{k=1}^{n} a_k = 96, \qquad \sum_{k=1}^{n} a_k^2 = 144, \qquad \sum_{k=1}^{n} a_k^3 = 216.$$

Problem 1.17.5 Demonstrate that for integer $n > 1$ we have,

$$n! < \left(\frac{n+1}{2}\right)^n.$$

Problem 1.17.6 Let $f(x) = (a+x)^5(a-x)^3$, $x \in [-a; a]$. Find the maximum value of de f using the AM-GM inequality.

Problem 1.17.7 Prove that the sequence $x_n = \left(1 + \dfrac{1}{n}\right)^n$, $n = 1, 2, \ldots$ is strictly increasing.

2 Differentiation

2.1 Some Topology

148 Definition Let $\mathbf{a} \in \mathbb{R}^n$ and let $\varepsilon > 0$. An *open ball* centred at $\mathbf{a}$ of radius ε is the set

$$B_\varepsilon(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x} - \mathbf{a}|| < \varepsilon\}.$$

An *open box* is a Cartesian product of open intervals

$$]a_1; b_1[\times]a_2; b_2[\times \cdots \times]a_{n-1}; b_{n-1}[\times]a_n; b_n[,$$

where the a_k, b_k are real numbers.

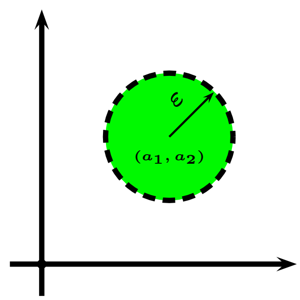

Figure 2.1: Open ball in $\mathbb{R}^2$.

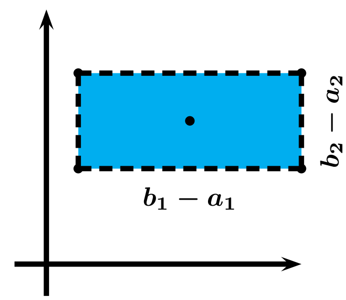

Figure 2.2: Open rectangle in $\mathbb{R}^2$.

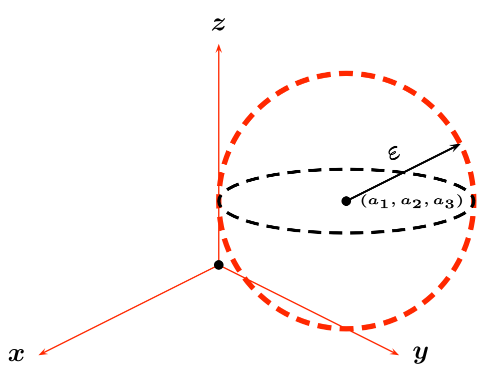

Figure 2.3: Open ball in $\mathbb{R}^3$.

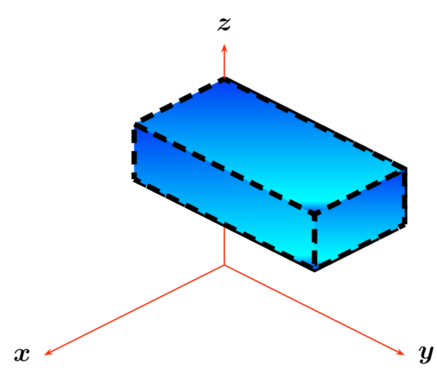

Figure 2.4: Open box in $\mathbb{R}^3$.

149 Example An open ball in $\mathbb{R}$ is an open interval, an open ball in $\mathbb{R}^2$ is an open disk (see figure 2.1) and an open ball in $\mathbb{R}^3$ is an open sphere (see figure 2.3). An open box in $\mathbb{R}$ is an open interval, an open box in $\mathbb{R}^2$ is a rectangle without its boundary (see figure 2.2) and an open box in $\mathbb{R}^3$ is a box without its boundary (see figure 2.4).

150 Definition A set $\mathscr{O} \subseteq \mathbb{R}^n$ is said to be *open* if for every point belonging to it we can surround the point by a sufficiently small open ball so that this balls lies completely within the set. That is, $\forall \mathrm{a} \in \mathscr{O} \ \ \exists \varepsilon > 0$ such that $B_\varepsilon(\mathrm{a}) \subseteq \mathscr{O}$.

151 Example The open interval $]-1;1[$ is open in $\mathbb{R}$. The interval $]-1;1]$ is not open, however, as no interval centred at 1 is totally contained in $]-1;1]$.

152 Example The region $]-1;1[\times]0;+\infty[$ is open in $\mathbb{R}^2$.

153 Example The ellipsoidal region $\left\{(x,y) \in \mathbb{R}^2 : x^2 + 4y^2 < 4\right\}$ is open in $\mathbb{R}^2$.

The reader will recognise that open boxes, open ellipsoids and their unions and finite intersections are open sets in $\mathbb{R}^n$.

154 Definition A set $\mathscr{F} \subseteq \mathbb{R}^n$ is said to be *closed* in $\mathbb{R}^n$ if its complement $\mathbb{R}^n \setminus \mathscr{F}$ is open.

155 Example The closed interval $[-1;1]$ is closed in $\mathbb{R}$, as its complement, $\mathbb{R}\setminus[-1;1] =]-\infty;-1[\cup]1;+\infty[$ is open in $\mathbb{R}$. The interval $]-1;1]$ is neither open nor closed in $\mathbb{R}$, however.

156 Example The region $[-1;1] \times [0;+\infty[\times[0;2]$ is closed in $\mathbb{R}^3$.

Homework

Problem 2.1.1 Determine whether the following subsets of $\mathbb{R}^2$ are open, closed, or neither, in $\mathbb{R}^2$.

1. $A = \{(x,y) \in \mathbb{R}^2 : |x| < 1, |y| < 1\}$
2. $B = \{(x,y) \in \mathbb{R}^2 : |x| < 1, |y| \leq 1\}$
3. $C = \{(x,y) \in \mathbb{R}^2 : |x| \leq 1, |y| \leq 1\}$
4. $D = \{(x,y) \in \mathbb{R}^2 : x^2 \leq y \leq x\}$
5. $E = \{(x,y) \in \mathbb{R}^2 : xy > 1\}$
6. $F = \{(x,y) \in \mathbb{R}^2 : xy \leq 1\}$
7. $G = \{(x,y) \in \mathbb{R}^2 : |y| \leq 9, x < y^2\}$

Problem 2.1.2 (Putnam Exam 1969) Let $p(x,y)$ be a polynomial with real coefficients in the real variables x and y, defined over the entire plane $\mathbb{R}^2$. What are the possibilities for the image (range) of $p(x,y)$?

Problem 2.1.3 (Putnam 1998) Let $\mathcal{F}$ be a finite collection of open disks in $\mathbb{R}^2$ whose union contains a set $E \subseteq \mathbb{R}^2$. Shew that there is a pairwise disjoint subcollection $D_k, k \geq 1$ in $\mathcal{F}$ such that

$$E \subseteq \bigcup_{j=1}^{n} 3D_j.$$

2.2 Multivariable Functions

Let $A \subseteq \mathbb{R}^n$. For most of this course, our concern will be functions of the form

$$f : A \to \mathbb{R}^m.$$

If $m = 1$, we say that f is a *scalar field*. If $m \geq 2$, we say that f is a *vector field*.

We would like to develop a calculus analogous to the situation in $\mathbb{R}$. In particular, we would like to examine limits, continuity, differentiability, and integrability of multivariable functions. Needless to say, the introduction of more variables greatly complicates the analysis. For example, recall that the graph of a function $f : A \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$. is the set

$$\{(\mathrm{x}, f(\mathrm{x})) : \mathrm{x} \in A)\} \subseteq \mathbb{R}^{n+m}.$$

If $m+n>3$, we have an object of more than three-dimensions! In the case $n=2, m=1$, we have a tri-dimensional surface. We will now briefly examine this case.

157 Definition Let $A \subseteq \mathbb{R}^2$ and let $f : A \to \mathbb{R}$ be a function. Given $c \in \mathbb{R}$, the *level curve* at $z=c$ is the curve resulting from the intersection of the surface $z=f(x,y)$ and the plane $z=c$, if there is such a curve.

158 Example The level curves of the surface $f(x,y)=x^2+y^2$ (an elliptic paraboloid) are the concentric circles

$$x^2+y^2=c, \qquad c>0.$$

Homework

Problem 2.2.1 Sketch the level curves for the following maps.

1. $(x,y) \mapsto x+y$
2. $(x,y) \mapsto xy$
3. $(x,y) \mapsto \min(|x|,|y|)$
4. $(x,y) \mapsto x^3-x$
5. $(x,y) \mapsto x^2+4y^2$
6. $(x,y) \mapsto \sin(x^2+y^2)$
7. $(x,y) \mapsto \cos(x^2-y^2)$

Problem 2.2.2 Sketch the level surfaces for the following maps.

1. $(x,y,z) \mapsto x+y+z$
2. $(x,y,z) \mapsto xyz$
3. $(x,y,z) \mapsto \min(|x|,|y|,|z|)$
4. $(x,y,z) \mapsto x^2+y^2$
5. $(x,y,z) \mapsto x^2+4y^2$
6. $(x,y,z) \mapsto \sin(z-x^2-y^2)$
7. $(x,y,z) \mapsto x^2+y^2+z^2$

2.3 Limits

We will start with the notion of *limit*.

159 Definition A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to have a limit $\mathbf{L} \in \mathbb{R}^m$ at $\mathbf{a} \in \mathbb{R}^n$ if $\forall \epsilon > 0 \exists \delta > 0$ such that

$$0 < ||\mathbf{x}-\mathbf{a}|| < \delta \implies ||f(\mathbf{x})-\mathbf{L}|| < \epsilon.$$

In such a case we write,

$$\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \mathbf{L}.$$

The notions of infinite limits, limits at infinity, and continuity at a point, are analogously defined. Limits in more than one dimension are perhaps trickier to find, as one must approach the test point from infinitely many directions.

160 Example Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2}$.

Solution: ▸ *We use the sandwich theorem. Observe that* $0 \leq x^2 \leq x^2+y^2$, *and so* $0 \leq \dfrac{x^2}{x^2+y^2} \leq 1$. *Thus*

$$\lim_{(x,y)\to(0,0)} 0 \leq \lim_{(x,y)\to(0,0)} \left|\frac{x^2y}{x^2+y^2}\right| \leq \lim_{(x,y)\to(0,0)} |y|,$$

and hence

$$\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2} = 0.$$

The Maple™ commands to graph this surface and find this limits appear below. Notice that Maple is unable to find the limit and so returns unevaluated.

```
> with(plots):
> plot3d(x^2*y/(x^2+y^2),x=-10..10,y=-10..10,axes=boxed,color=x^2+y^2);
> limit(x^2*y/(x^2+y^2),x=0,y=0);
```

◀

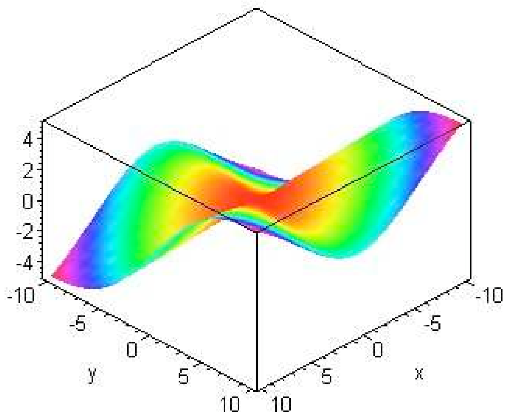

Figure 2.5: $(x, y) \mapsto \dfrac{x^2 y}{x^2 + y^2}$.

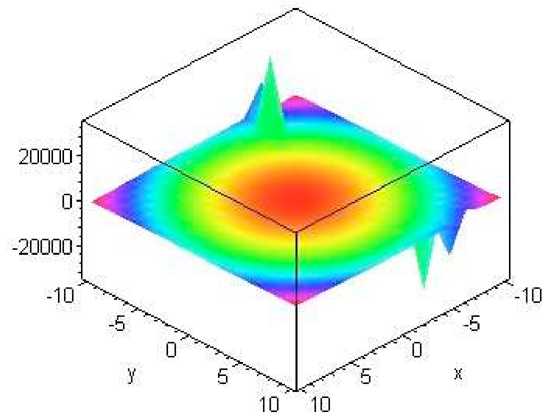

Figure 2.6: $(x, y) \mapsto \dfrac{x^5 y^3}{x^6 + y^4}$.

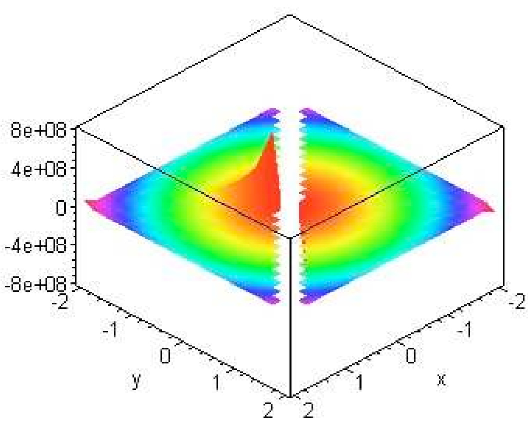

Figure 2.7: Example 162.

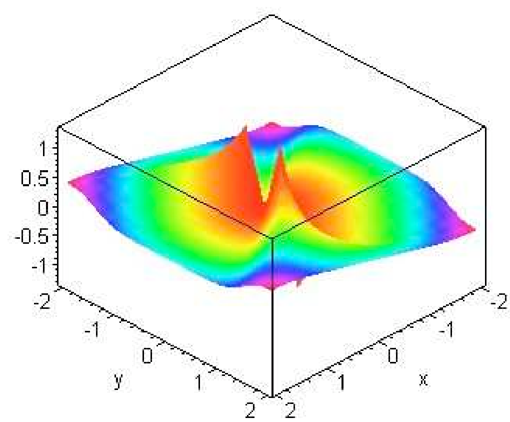

Figure 2.8: Example 163.

161 Example Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^5 y^3}{x^6 + y^4}$.

Solution: ► *Either $|x| \leq |y|$ or $|x| \geq |y|$. Observe that if $|x| \leq |y|$, then*

$$\left|\frac{x^5y^3}{x^6+y^4}\right| \leq \frac{y^8}{y^4} = y^4.$$

If $|y| \leq |x|$, then

$$\left|\frac{x^5y^3}{x^6+y^4}\right| \leq \frac{x^8}{x^6} = x^2.$$

Thus

$$\left|\frac{x^5y^3}{x^6+y^4}\right| \leq \max(y^4, x^2) \leq y^4 + x^2 \longrightarrow 0,$$

as $(x, y) \to (0, 0)$.
Aliter: *Let $X = x^3, Y = y^2$.*

$$\left|\frac{x^5y^3}{x^6+y^4}\right| = \frac{X^{5/3}Y^{3/2}}{X^2+Y^2}.$$

Passing to polar coordinates $X = \rho\cos\theta, Y = \rho\sin\theta$, we obtain

$$\left|\frac{x^5y^3}{x^6+y^4}\right| = \frac{X^{5/3}Y^{3/2}}{X^2+Y^2} = \rho^{5/3+3/2-2}|\cos\theta|^{5/3}|\sin\theta|^{3/2} \leq \rho^{7/6} \to 0,$$

as $(x, y) \to (0, 0)$. ◄

162 Example Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{1+x+y}{x^2-y^2}$.

Solution: ► *When $y = 0$,*

$$\frac{1+x}{x^2} \to +\infty,$$

as $x \to 0$. When $x = 0$,

$$\frac{1+y}{-y^2} \to -\infty,$$

as $y \to 0$. The limit does not exist. ◄

163 Example Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy^6}{x^6+y^8}$.

Solution: ► *Putting $x = t^4, y = t^3$, we find*

$$\frac{xy^6}{x^6+y^8} = \frac{1}{2t^2} \to +\infty,$$

as $t \to 0$. But when $y = 0$, the function is 0. Thus the limit does not exist. ◄

164 Example Find $\displaystyle\lim_{(x,y)\to(0,0)} \frac{((x-1)^2+y^2)\log_e((x-1)^2+y^2)}{|x|+|y|}$.

Solution: ► *When $y = 0$ we have*

$$\frac{2(x-1)^2\ln(|1-x|)}{|x|} \sim -\frac{2x}{|x|},$$

and so the function does not have a limit at $(0, 0)$. ◄

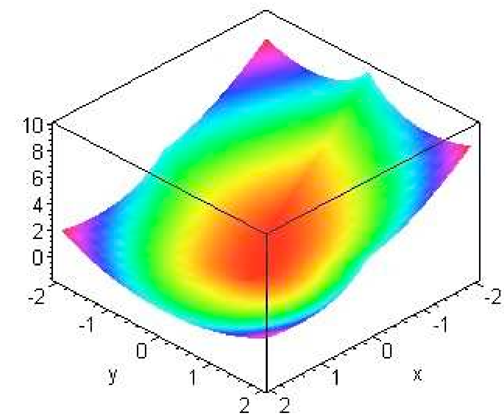

Figure 2.9: Example 164.

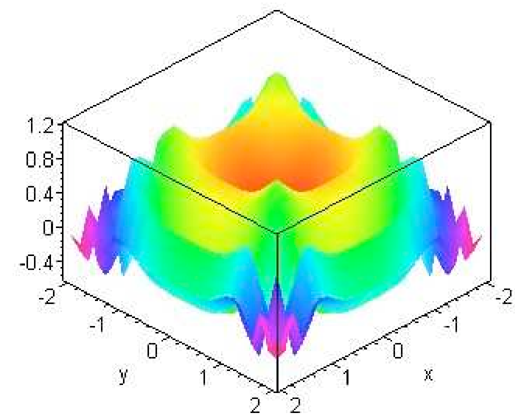

Figure 2.10: Example 165.

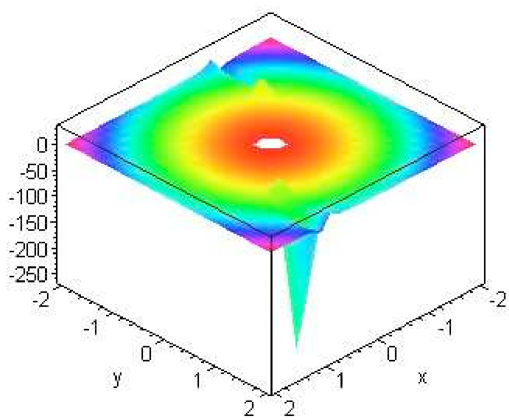

Figure 2.11: Example 166.

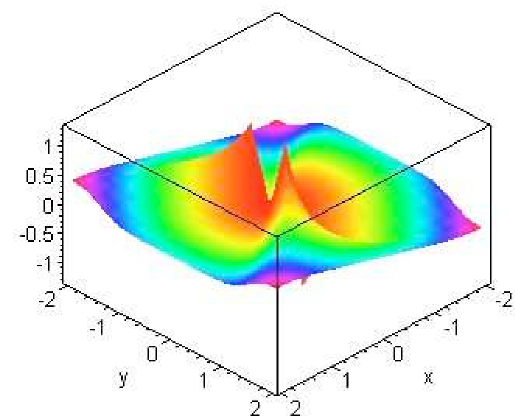

Figure 2.12: Example 163.

165 Example Find $\lim_{(x,y)\to(0,0)} \dfrac{\sin(x^4)+\sin(y^4)}{\sqrt{x^4+y^4}}$.

Solution: ► $\sin(x^4)+\sin(y^4) \leq x^4+y^4$ *and so*

$$\left|\frac{\sin(x^4)+\sin(y^4)}{\sqrt{x^4+y^4}}\right| \leq \sqrt{x^4+y^4} \to 0,$$

as $(x,y) \to (0,0)$. ◄

166 Example Find $\lim_{(x,y)\to(0,0)} \dfrac{\sin x - y}{x - \sin y}$.

Solution: ► *When* $y = 0$ *we obtain*

$$\frac{\sin x}{x} \to 1,$$

as $x \to 0$. *When* $y = x$ *the function is identically* -1. *Thus the limit does not exist.* ◄

If $f : \mathbb{R}^2 \to \mathbb{R}$, it may be that the limits

$$\lim_{y\to y_0}\left(\lim_{x\to x_0} f(x,y)\right), \qquad \lim_{x\to x_0}\left(\lim_{y\to y_0} f(x,y)\right),$$

both exist. These are called the *iterated limits of* f *as* $(x,y) \to (x_0,y_0)$. The following possibilities might occur.

1. If $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ exists, then each of the iterated limits $\lim_{y\to y_0}\left(\lim_{x\to x_0} f(x,y)\right)$ and $\lim_{x\to x_0}\left(\lim_{y\to y_0} f(x,y)\right)$ exists.

2. If the iterated limits exist and $\lim_{y\to y_0}\left(\lim_{x\to x_0} f(x,y)\right) \neq \lim_{x\to x_0}\left(\lim_{y\to y_0} f(x,y)\right)$ then $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ does not exist.

3. It may occur that $\lim_{y\to y_0}\left(\lim_{x\to x_0} f(x,y)\right) = \lim_{x\to x_0}\left(\lim_{y\to y_0} f(x,y)\right)$, but that $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ does not exist.

4. It may occur that $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ exists, but one of the iterated limits does not.

Homework

Problem 2.3.1 Sketch the domain of definition of $(x,y) \mapsto \sqrt{4-x^2-y^2}$.

Problem 2.3.2 Sketch the domain of definition of $(x,y) \mapsto \log(x+y)$.

Problem 2.3.3 Sketch the domain of definition of $(x,y) \mapsto \dfrac{1}{x^2+y^2}$.

Problem 2.3.4 Find $\lim_{(x,y)\to(0,0)} (x^2+y^2)\sin\dfrac{1}{xy}$.

Problem 2.3.5 Find $\lim_{(x,y)\to(0,2)} \dfrac{\sin xy}{x}$.

Problem 2.3.6 For what c will the function

$$f(x,y) = \begin{cases} \sqrt{1-x^2-4y^2}, & \text{if } x^2+4y^2 \leq 1, \\ c, & \text{if } x^2+4y^2 > 1 \end{cases}$$

be continuous everywhere on the xy-plane?

Problem 2.3.7 Find

$$\lim_{(x,y)\to(0,0)} \sqrt{x^2+y^2}\sin\frac{1}{x^2+y^2}.$$

Problem 2.3.8 Find

$$\lim_{(x,y)\to(+\infty,+\infty)} \frac{\max(|x|,|y|)}{\sqrt{x^4+y^4}}.$$

Problem 2.3.9 Find

$$\lim_{(x,y)\to(0,0} \frac{2x^2\sin y^2 + y^4 e^{-|x|}}{\sqrt{x^2+y^2}}.$$

Problem 2.3.10 Demonstrate that

$$\lim_{(x,y,z)\to(0,0,0)} \frac{x^2y^2z^2}{x^2+y^2+z^2} = 0.$$

Problem 2.3.11 Prove that

$$\lim_{x\to 0}\left(\lim_{y\to 0}\frac{x-y}{x+y}\right) = 1 = -\lim_{y\to 0}\left(\lim_{x\to 0}\frac{x-y}{x+y}\right).$$

Does $\lim_{(x,y)\to(0,0)} \frac{x-y}{x+y}$ exist?.

Problem 2.3.12 Let

$$f(x,y) = \begin{cases} x\sin\frac{1}{x} + y\sin\frac{1}{y} & \text{if } x\neq 0,\ y\neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Prove that $\lim_{(x,y)\to(0,0)} f(x,y)$ exists, but that the iterated limits $\lim_{x\to 0}\left(\lim_{y\to 0} f(x,y)\right)$ and $\lim_{y\to 0}\left(\lim_{x\to 0} f(x,y)\right)$ do not exist.

Problem 2.3.13 Prove that

$$\lim_{x\to 0}\left(\lim_{y\to 0}\frac{x^2y^2}{x^2y^2+(x-y)^2}\right) = 0,$$

and that

$$\lim_{y\to 0}\left(\lim_{x\to 0}\frac{x^2y^2}{x^2y^2+(x-y)^2}\right) = 0,$$

but still $\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2y^2+(x-y)^2}$ does not exist.

2.4 Definition of the Derivative

Before we begin, let us introduce some necessary notation. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We write $f(h) = \mathrm{o}(h)$ if $f(h)$ goes faster to 0 than h, that is, if $\lim_{h\to 0}\frac{f(h)}{h} = 0$. For example, $h^3 + 2h^2 = \mathrm{o}(h)$, since

$$\lim_{h\to 0}\frac{h^3+2h^2}{h} = \lim_{h\to 0} h^2 + 2h = 0.$$

We now define the derivative in the multidimensional space $\mathbb{R}^n$. Recall that in one variable, a function $g : \mathbb{R}\to\mathbb{R}$ is said to be differentiable at $x = a$ if the limit

$$\lim_{x\to a}\frac{g(x)-g(a)}{x-a} = g'(a)$$

exists. The limit condition above is equivalent to saying that

$$\lim_{x\to a}\frac{g(x)-g(a)-g'(a)(x-a)}{x-a} = 0,$$

or equivalently,

$$\lim_{h\to 0}\frac{g(a+h)-g(a)-g'(a)(h)}{h} = 0.$$

We may write this as

$$g(a+h)-g(a) = g'(a)(h) + \mathrm{o}(h).$$

The above analysis provides an analogue definition for the higher-dimensional case. Observe that since we may not divide by vectors, the corresponding definition in higher dimensions involves quotients of norms.

167 Definition Let $A \subseteq \mathbb{R}^n$. A function $f : A \to \mathbb{R}^m$ is said to be *differentiable* at $\mathbf{a}\in A$ if there is a linear transformation, called the *derivative of f at* $\mathbf{a}$, $\mathscr{D}_{\mathbf{a}}(f) : \mathbb{R}^n\to\mathbb{R}^m$ such that

$$\lim_{\mathbf{x}\to\mathbf{a}}\frac{||f(\mathbf{x})-f(\mathbf{a})-\mathscr{D}_{\mathbf{a}}(f)(\mathbf{x}-\mathbf{a})||}{||\mathbf{x}-\mathbf{a}||} = 0.$$

Equivalently, f is differentiable at $\mathbf{a}$ if there is a linear transformation $\mathscr{D}_{\mathbf{a}}(f)$ such that

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \mathscr{D}_{\mathbf{a}}(f)(\mathbf{h}) + \mathrm{o}\left(||\mathbf{h}||\right),$$

as $\mathbf{h} \to \mathbf{0}$.

☞ *The condition for differentiability at* $\mathbf{a}$ *is equivalent to*

$$f(\mathbf{x}) - f(\mathbf{a}) = \mathscr{D}_{\mathbf{a}}(f)(\mathbf{x}-\mathbf{a}) + \mathrm{o}\left(||\mathbf{x}-\mathbf{a}||\right),$$

as $\mathbf{x} \to \mathbf{a}$.

168 Theorem If A is an open set in definition 167, $\mathscr{D}_{\mathbf{a}}(f)$ is uniquely determined.

Proof: *Let* $L : \mathbb{R}^n \to \mathbb{R}^m$ *be another linear transformation satisfying definition 167. We must prove that* $\forall \mathbf{v} \in \mathbb{R}^n, L(\mathbf{v}) = \mathscr{D}_{\mathbf{a}}(f)(\mathbf{v})$. *Since* A *is open,* $\mathbf{a}+\mathbf{h} \in A$ *for sufficiently small* $||\mathbf{h}||$. *By definition, as* $\mathbf{h} \to \mathbf{0}$, *we have*

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \mathscr{D}_{\mathbf{a}}(f)(\mathbf{h}) + \mathrm{o}\left(||\mathbf{h}||\right).$$

and

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = L(\mathbf{h}) + \mathrm{o}\left(||\mathbf{h}||\right).$$

Now, observe that

$$\mathscr{D}_{\mathbf{a}}(f)(\mathbf{v}) - L(\mathbf{v}) = \mathscr{D}_{\mathbf{a}}(f)(\mathbf{h}) - f(\mathbf{a}+\mathbf{h}) + f(\mathbf{a}) + f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h}).$$

By the triangle inequality,

$$\begin{aligned} ||\mathscr{D}_{\mathbf{a}}(f)(\mathbf{v}) - L(\mathbf{v})|| &\leq ||\mathscr{D}_{\mathbf{a}}(f)(\mathbf{h}) - f(\mathbf{a}+\mathbf{h}) + f(\mathbf{a})|| \\ &\quad + ||f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h})|| \\ &= \mathrm{o}\left(||\mathbf{h}||\right) + \mathrm{o}\left(||\mathbf{h}||\right) \\ &= \mathrm{o}\left(||\mathbf{h}||\right), \end{aligned}$$

as $\mathbf{h} \to \mathbf{0}$ *This means that*

$$||L(\mathbf{v}) - \mathscr{D}_{\mathbf{a}}(f)(\mathbf{v})|| \to 0,$$

i.e., $L(\mathbf{v}) = \mathscr{D}_{\mathbf{a}}(f)(\mathbf{v})$, *completing the proof.* ❑

☞ *If* $A = \{\mathbf{a}\}$, *a singleton, then* $\mathscr{D}_{\mathbf{a}}(f)$ *is not uniquely determined. For* $||\mathbf{x}-\mathbf{a}|| < \delta$ *holds only for* $\mathbf{x} = \mathbf{a}$, *and so* $f(\mathbf{x}) = f(\mathbf{a})$. *Any linear transformation* T *will satisfy the definition, as* $T(\mathbf{x}-\mathbf{a}) = T(\mathbf{0}) = \mathbf{0}$, *and*

$$||f(\mathbf{x}) - f(\mathbf{a}) - T(\mathbf{x}-\mathbf{a})|| = ||\mathbf{0}|| = 0,$$

identically.

169 Example If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then $\mathscr{D}_{\mathbf{a}}(L) = L$, for any $\mathbf{a} \in \mathbb{R}^n$.

Solution: ▶ *Since* $\mathbb{R}^n$ *is an open set, we know that* $\mathscr{D}_{\mathbf{a}}(L)$ *uniquely determined. Thus if* L *satisfies definition 167, then the claim is established. But by linearity*

$$||L(\mathbf{x}) - L(\mathbf{a}) - L(\mathbf{x}-\mathbf{a})|| = ||L(\mathbf{x}) - L(\mathbf{a}) - L(\mathbf{x}) + L(\mathbf{a})|| = ||\mathbf{0}|| = 0,$$

whence the claim follows. ◀

170 Example Let

$$f : \begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R} \\ (\overrightarrow{x}, \overrightarrow{y}) & \mapsto & \overrightarrow{x} \bullet \overrightarrow{y} \end{array}$$

be the usual dot product in $\mathbb{R}^3$. Shew that f is differentiable and that

$$\mathscr{D}_{(\overrightarrow{x}, \overrightarrow{y})} f(\overrightarrow{h}, \overrightarrow{k}) = \overrightarrow{x} \bullet \overrightarrow{k} + \overrightarrow{h} \bullet \overrightarrow{y}.$$

Solution: ► *We have*

$$\begin{aligned} f(\overrightarrow{x} + \overrightarrow{h}, \overrightarrow{y} + \overrightarrow{k}) - f(\overrightarrow{x}, \overrightarrow{y}) &= (\overrightarrow{x} + \overrightarrow{h})\bullet(\overrightarrow{y} + \overrightarrow{k}) - \overrightarrow{x}\bullet\overrightarrow{y} \\ &= \overrightarrow{x}\bullet\overrightarrow{y} + \overrightarrow{x}\bullet\overrightarrow{k} + \overrightarrow{h}\bullet\overrightarrow{y} + \overrightarrow{h}\bullet\overrightarrow{k} - \overrightarrow{x}\bullet\overrightarrow{y} \\ &= \overrightarrow{x}\bullet\overrightarrow{k} + \overrightarrow{h}\bullet\overrightarrow{y} + \overrightarrow{h}\bullet\overrightarrow{k}. \end{aligned}$$

As $(\overrightarrow{h}, \overrightarrow{k}) \to (\overrightarrow{0}, \overrightarrow{0})$, *we have by the Cauchy-Buniakovskii-Schwarz inequality,* $|\overrightarrow{h}\bullet\overrightarrow{k}| \leq \left|\left|\overrightarrow{h}\right|\right|\left|\left|\overrightarrow{k}\right|\right| = o\left(\left|\left|\overrightarrow{h}\right|\right|\right)$, *which proves the assertion.* ◄

Just like in the one variable case, differentiability at a point, implies continuity at that point.

171 Theorem Suppose $A \subseteq \mathbb{R}^n$ is open and $f : A \to \mathbb{R}^n$ is differentiable on A. Then f is continuous on A.

Proof: *Given* $\mathrm{a} \in A$, *we must shew that*

$$\lim_{\mathrm{x}\to\mathrm{a}} f(\mathrm{x}) = f(\mathrm{a}).$$

Since f *is differentiable at* a *we have*

$$f(\mathrm{x}) - f(\mathrm{a}) = \mathscr{D}_{\mathrm{a}}(f)(\mathrm{x} - \mathrm{a}) + o(||\mathrm{x} - \mathrm{a}||),$$

and so

$$f(\mathrm{x}) - f(\mathrm{a}) \to 0,$$

as $\mathrm{x} \to \mathrm{a}$, *proving the theorem.* ❑

Homework

Problem 2.4.1 Let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation and

$$F : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^3 \\ \overrightarrow{x} & \mapsto & \overrightarrow{x} \times L(\overrightarrow{x}) \end{array}.$$

Shew that F is differentiable and that

$$\mathscr{D}_{\mathrm{x}}(F)(\overrightarrow{h}) = \overrightarrow{x} \times L(\overrightarrow{h}) + \overrightarrow{h} \times L(\overrightarrow{x}).$$

Problem 2.4.2 Let $f : \mathbb{R}^n \to \mathbb{R}, n \geq 1, f(\overrightarrow{x}) = \left|\left|\overrightarrow{x}\right|\right|$ be the usual norm in $\mathbb{R}^n$, with $\left|\left|\overrightarrow{x}\right|\right|^2 = \overrightarrow{x}\bullet\overrightarrow{x}$. Prove that

$$\mathscr{D}_{\mathrm{x}}(f)(\overrightarrow{v}) = \frac{\overrightarrow{x}\bullet\overrightarrow{v}}{\left|\left|\overrightarrow{x}\right|\right|},$$

for $\overrightarrow{x} \neq \overrightarrow{0}$, but that f is not differentiable at $\overrightarrow{0}$.

2.5 The Jacobi Matrix

We now establish a way which simplifies the process of finding the derivative of a function at a given point.

172 Definition Let $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}^m$, and put

$$f(\mathrm{x}) = \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{bmatrix}.$$

Here $f_i : \mathbb{R}^n \to \mathbb{R}$. The *partial derivative* $\dfrac{\partial f_i}{\partial x_j}(\mathrm{x})$ is defined as

$$\frac{\partial f_i}{\partial x_j}(\mathrm{x}) = \lim_{h\to 0} \frac{f_i(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f_i(x_1, x_2, \ldots, x_j, \ldots, x_n)}{h},$$

whenever this limit exists.

To find partial derivatives with respect to the j-th variable, we simply keep the other variables fixed and differentiate with respect to the j-th variable.

173 Example If $f : \mathbb{R}^3 \to \mathbb{R}$, and $f(x, y, z) = x + y^2 + z^3 + 3xy^2z^3$ then

$$\frac{\partial f}{\partial x}(x, y, z) = 1 + 3y^2z^3,$$

$$\frac{\partial f}{\partial y}(x, y, z) = 2y + 6xyz^3,$$

and

$$\frac{\partial f}{\partial z}(x, y, z) = 3z^2 + 9xy^2z^2.$$

The Maple™ commands to find these follow.

```
>  f:=(x,y,z)->x+y^2+z^3+3*x*y^2*z^3;
>  diff(f(x,y,z),x);
>  diff(f(x,y,z),y);
>  diff(f(x,y,z),z);
```

Since the derivative of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, it can be represented by aid of matrices. The following theorem will allow us to determine the matrix representation for $\mathscr{D}_{\mathrm{a}}(f)$ under the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$.

174 Theorem Let

$$f(\mathrm{x}) = \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{bmatrix}.$$

Suppose $A \subseteq \mathbb{R}^n$ is an open set and $f : A \to \mathbb{R}^m$ is differentiable. Then each partial derivative $\dfrac{\partial f_i}{\partial x_j}(\mathrm{x})$ exists, and the matrix representation of $\mathscr{D}_{\mathrm{x}}(f)$ with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ is the *Jacobi matrix*

$$f'(\mathrm{x}) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathrm{x}) & \dfrac{\partial f_1}{\partial x_2}(\mathrm{x}) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\mathrm{x}) \\ \dfrac{\partial f_2}{\partial x_1}(\mathrm{x}) & \dfrac{\partial f_2}{\partial x_2}(\mathrm{x}) & \cdots & \dfrac{\partial f_2}{\partial x_n}(\mathrm{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(\mathrm{x}) & \dfrac{\partial f_n}{\partial x_2}(\mathrm{x}) & \cdots & \dfrac{\partial f_n}{\partial x_n}(\mathrm{x}) \end{bmatrix}.$$

Proof: *Let $\overrightarrow{\mathrm{e}}_j, 1 \leq j \leq n$, be the standard basis for $\mathbb{R}^n$. To obtain the Jacobi matrix, we must compute $\mathscr{D}_{\mathrm{x}}(f)(\overrightarrow{\mathrm{e}}_j)$, which will give us the j-th column of the Jacobi matrix. Let $f'(\mathrm{x}) = (J_{ij})$, and observe that*

$$\mathscr{D}_{\mathrm{x}}(f)(\overrightarrow{\mathrm{e}}_j) = \begin{bmatrix} J_{1j} \\ J_{2j} \\ \vdots \\ J_{nj} \end{bmatrix}.$$

and put $\mathrm{y} = \mathrm{x} + \varepsilon \overrightarrow{\mathrm{e}}_j, \varepsilon \in \mathbb{R}$. Notice that

$$\frac{||f(\mathrm{y}) - f(\mathrm{x}) - \mathscr{D}_{\mathrm{x}}(f)(\mathrm{y} - \mathrm{x})||}{||\mathrm{y} - \mathrm{x}||}$$
$$= \frac{||f(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f(x_1, x_2, \ldots, x_j, \ldots, x_n) - \varepsilon\mathscr{D}_{\mathrm{x}}(f)(\overrightarrow{\mathrm{e}}_j)||}{|\varepsilon|}.$$

Since the sinistral side $\to 0$ as $\varepsilon \to 0$, the so does the i-th component of the numerator, and so,

$$\frac{|f_i(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f_i(x_1, x_2, \ldots, x_j, \ldots, x_n) - \varepsilon J_{ij}|}{|\varepsilon|} \to 0.$$

This entails that

$$J_{ij} = \lim_{\varepsilon \to 0} \frac{f_i(x_1, x_2, \ldots, x_j + \varepsilon, \ldots, x_n) - f_i(x_1, x_2, \ldots, x_j, \ldots, x_n)}{\varepsilon} = \frac{\partial f_i}{\partial x_j}(\mathrm{x}).$$

This finishes the proof. ❑

☞ *Strictly speaking, the Jacobi matrix is not the derivative of a function at a point. It is a matrix representation of the derivative in the standard basis of $\mathbb{R}^n$. We will abuse language, however, and refer to f' when we mean the Jacobi matrix of f.*

175 Example Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$f(x, y) = (xy + yz, \log_e xy).$$

Compute the Jacobi matrix of f.

Solution: ► *The Jacobi matrix is the 2×3 matrix*

$$f'(x, y) = \begin{bmatrix} \frac{\partial f_1}{\partial x}(x, y) & \frac{\partial f_1}{\partial y}(x, y) & \frac{\partial f_1}{\partial z}(x, y) \\ \frac{\partial f_2}{\partial x}(x, y) & \frac{\partial f_2}{\partial y}(x, y) & \frac{\partial f_2}{\partial z}(x, y) \end{bmatrix} = \begin{bmatrix} y & x + z & y \\ \frac{1}{x} & \frac{1}{y} & 0 \end{bmatrix}.$$

◄

176 Example Let $f(\rho, \theta, z) = (\rho \cos\theta, \rho \sin\theta, z)$ be the function which changes from cylindrical coordinates to Cartesian coordinates. We have

$$f'(\rho, \theta, z) = \begin{bmatrix} \cos\theta & -\rho\sin\theta & 0 \\ \sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

177 Example Let $f(\rho, \phi, \theta) = (\rho\cos\theta\sin\phi, \rho\sin\theta\sin\phi, \rho\cos\phi)$ be the function which changes from spherical coordinates to Cartesian coordinates. We have

$$f'(\rho, \phi, \theta) = \begin{bmatrix} \cos\theta\sin\phi & \rho\cos\theta\cos\phi & -\rho\sin\phi\sin\theta \\ \sin\theta\sin\phi & \rho\sin\theta\cos\phi & \rho\cos\theta\sin\phi \\ \cos\phi & -\rho\sin\phi & 0 \end{bmatrix}.$$

The Jacobi matrix provides a convenient computational tool to compute the derivative of a function at a point. Thus differentiability at a point implies that the partial derivatives of the function exist at the point. The converse, however, is not true.

178 Example Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x, y) = \begin{cases} y & \text{if} \quad x = 0, \\ x & \text{if} \quad y = 0, \\ 1 & \text{if} \quad xy \neq 0. \end{cases}$$

Observe that f is not continuous at $(0,0)$ ($f(0,0) = 0$ but $f(x,y) = 1$ for values arbitrarily close to $(0,0)$), and hence, it is not differentiable there. We have however, $\dfrac{\partial f}{\partial x}(0,0) = \dfrac{\partial f}{\partial y}(0,0) = 1$. Thus even if both partial derivatives exist at $(0,0)$ is no guarantee that the function will be differentiable at $(0,0)$. You should also notice that both partial derivatives are not continuous at $(0,0)$.

We have, however, the following.

179 Theorem Let $A \subseteq \mathbb{R}^n$ be an open set, and let $f : \mathbb{R}^n \to \mathbb{R}^m$. Put $f = \begin{bmatrix} f_1 \\ f_2 \\ \cdots \\ f_m \end{bmatrix}$. If each of the partial derivatives $\mathscr{D}_j f_i$ exists and is continuous on A, then f is differentiable on A.

The concept of *repeated partial derivatives* is akin to the concept of repeated differentiation. Similarly with the concept of implicit partial differentiation. The following examples should be self-explanatory.

180 Example Let $f(u,v,w) = e^u v \cos w$. Determine $\dfrac{\partial^2}{\partial u \partial v} f(u,v,w)$ at $(1, -1, \dfrac{\pi}{4})$.

Solution: ► *We have*

$$\frac{\partial^2}{\partial u \partial v}(e^u v \cos w) = \frac{\partial}{\partial u}(e^u \cos w) = e^u \cos w,$$

which is $\dfrac{e\sqrt{2}}{2}$ *at the desired point.* ◄

181 Example The equation $z^{xy} + (xy)^z + xy^2z^3 = 3$ defines z as an implicit function of x and y. Find $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ at $(1,1,1)$.

Solution: ► *We have*

$$\begin{aligned} \frac{\partial}{\partial x} z^{xy} &= \frac{\partial}{\partial x} e^{xy \log z} \\ &= \left(y \log z + \frac{xy}{z} \frac{\partial z}{\partial x} \right) z^{xy}, \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial x} (xy)^z &= \frac{\partial}{\partial x} e^{z \log xy} \\ &= \left(\frac{\partial z}{\partial x} \log xy + \frac{z}{x} \right) (xy)^z, \end{aligned}$$

$$\frac{\partial}{\partial x} xy^2z^3 = y^2z^3 + 3xy^2z^2 \frac{\partial z}{\partial x},$$

Hence, at $(1,1,1)$ *we have*

$$\frac{\partial z}{\partial x} + 1 + 1 + 3\frac{\partial z}{\partial x} = 0 \implies \frac{\partial z}{\partial x} = -\frac{1}{2}.$$

Similarly,

$$\begin{aligned} \frac{\partial}{\partial y} z^{xy} &= \frac{\partial}{\partial y} e^{xy \log z} \\ &= \left(x \log z + \frac{xy}{z} \frac{\partial z}{\partial y} \right) z^{xy}, \end{aligned}$$

$$
\begin{aligned}
\frac{\partial}{\partial y}(xy)^z &= \frac{\partial}{\partial y} e^{z\log xy} \\
&= \left(\frac{\partial z}{\partial y}\log xy + \frac{z}{y}\right)(xy)^z,
\end{aligned}
$$

$$
\frac{\partial}{\partial y} xy^2z^3 = 2xyz^3 + 3xy^2z^2\frac{\partial z}{\partial y},
$$

Hence, at $(1,1,1)$ we have

$$
\frac{\partial z}{\partial y} + 1 + 2 + 3\frac{\partial z}{\partial y} = 0 \implies \frac{\partial z}{\partial y} = -\frac{3}{4}.
$$

◄

Just like in the one-variable case, we have the following rules of differentiation. Let $A \subseteq \mathbb{R}^n, B \subseteq \mathbb{R}^m$ be open sets $f, g : A \to \mathbb{R}^m, \alpha \in \mathbb{R}$, be differentiable on A, $h : B \to \mathbb{R}^l$ be differentiable on B, and $f(A) \subseteq B$. Then we have

- **Addition Rule:** $\mathscr{D}_{\mathrm{x}}((f + \alpha g)) = \mathscr{D}_{\mathrm{x}}(f) + \alpha\mathscr{D}_{\mathrm{x}}(g)$.
- **Chain Rule:** $\mathscr{D}_{\mathrm{x}}((h \circ f)) = \left(\mathscr{D}_{f(\mathrm{x})}(h)\right) \circ (\mathscr{D}_{\mathrm{x}}(f))$.

Since composition of linear mappings expressed as matrices is matrix multiplication, the Chain Rule takes the alternative form when applied to the Jacobi matrix.

$$
(h \circ f)' = (h' \circ f)(f'). \tag{2.1}
$$

182 Example Let

$$
f(u, v) = \begin{bmatrix} ue^v \\ u + v \\ uv \end{bmatrix},
$$

$$
h(x, y) = \begin{bmatrix} x^2 + y \\ y + z \end{bmatrix}.
$$

Find $(f \circ h)'(x, y)$.

Solution: ► *We have*

$$
f'(u, v) = \begin{bmatrix} e^v & ue^v \\ 1 & 1 \\ v & u \end{bmatrix},
$$

and

$$
h'(x, y) = \begin{bmatrix} 2x & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.
$$

Observe also that

$$
f'(h(x, y)) = \begin{bmatrix} e^{y+z} & (x^2 + y)e^{y+z} \\ 1 & 1 \\ y + z & x^2 + y \end{bmatrix}.
$$

Hence

$$\begin{aligned}(f\circ h)'(x,y) &= f'(h(x,y))h'(x,y)\\ &= \begin{bmatrix} e^{y+z} & (x^2+y)e^{y+z}\\ 1 & 1\\ y+z & x^2+y\end{bmatrix}\begin{bmatrix}2x & 1 & 0\\ 0 & 1 & 1\end{bmatrix}\\ &= \begin{bmatrix} 2xe^{y+z} & (1+x^2+y)e^{y+z} & (x^2+y)e^{y+z}\\ 2x & 2 & 1\\ 2xy+2xz & x^2+2y+z & x^2+y\end{bmatrix}.\end{aligned}$$

◀

183 Example Let

$$f:\mathbb{R}^2\to\mathbb{R},\quad f(u,v)=u^2+e^v,$$
$$u,v:\mathbb{R}^3\to\mathbb{R}\quad u(x,y)=xz,\ v(x,y)=y+z.$$

Put $h(x,y)=f\begin{bmatrix}u(x,y,z)\\ v(x,y,z)\end{bmatrix}$. Find the partial derivatives of h.

Solution: ▶ *Put* $g:\mathbb{R}^3\to\mathbb{R}^2, g(x,y)=\begin{bmatrix}u(x,y)\\ v(x,y\end{bmatrix}=\begin{bmatrix}xz\\ y+z\end{bmatrix}$. *Observe that* $h=f\circ g$. *Now,*

$$g'(x,y)=\begin{bmatrix}z & 0 & x\\ 0 & 1 & 1\end{bmatrix},$$

$$f'(u,v)=\begin{bmatrix}2u & e^v\end{bmatrix},$$

$$f'(h(x,y))=\begin{bmatrix}2xz & e^{y+z}\end{bmatrix}.$$

Thus

$$\begin{aligned}\begin{bmatrix}\frac{\partial h}{\partial x}(x,y) & \frac{\partial h}{\partial y}(x,y) & \frac{\partial h}{\partial z}(x,y)\end{bmatrix} &= h'(x,y)\\ &= (f'(g(x,y)))(g'(x,y))\\ &= \begin{bmatrix}2xz & e^{y+z}\end{bmatrix}\begin{bmatrix}z & 0 & x\\ 0 & 1 & 1\end{bmatrix}\\ &= \begin{bmatrix}2xz^2 & e^{y+z} & 2x^2z+e^{y+z}\end{bmatrix}\end{aligned}.$$

Equating components, we obtain

$$\frac{\partial h}{\partial x}(x,y)=2xz^2,$$
$$\frac{\partial h}{\partial y}(x,y)=e^{y+z},$$
$$\frac{\partial h}{\partial z}(x,y)=2x^2z+e^{y+z}.$$

◀

Under certain conditions we may differentiate under the integral sign.

184 Theorem (Differentiation under the integral sign) Let $f\colon [a,b]\times Y\to\mathbb{R}$ be a function, with $[a,b]$ being a closed interval, and Y being a closed and bounded subset of $\mathbb{R}$. Suppose that both $f(x,y)$ and $\dfrac{\partial}{\partial x}f(x,y)$ are continuous in the variables x and y jointly. Then $\int_Y f(x.y)\,dy$ exists as a continuously differentiable function of x on $[a,b]$, with derivative

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_Y f(x,y)\,\mathrm{d}y=\int_Y\frac{\partial}{\partial x}f(x,y)\,\mathrm{d}y.$$

185 Example Prove that

$$F(x)=\int_0^{\pi/2}\log(\sin^2\theta+x^2\cos^2\theta)\mathrm{d}\theta=\pi\log\frac{x+1}{2}.$$

Solution: ► *Differentiating under the integral,*

$$\begin{aligned}F'(x)&=\int_0^{\pi/2}\frac{\partial}{\partial x}\log(\sin^2\theta+x^2\cos^2\theta)\mathrm{d}\theta\\&=2x\int_0^{\pi/2}\frac{\cos^2\theta}{\sin^2\theta+x^2\cos^2\theta}\mathrm{d}\theta\end{aligned}$$

. The above implies that

$$\begin{aligned}\frac{(x^2-1)}{2x}\cdot F'(x)&=\int_0^{\pi/2}\frac{(x^2-1)\cos^2\theta}{\sin^2\theta+x^2\cos^2\theta}\mathrm{d}\theta\\&=\int_0^{\pi/2}\frac{x^2\cos^2\theta+\sin^2\theta-1}{\sin^2\theta+x^2\cos^2\theta}\mathrm{d}\theta\\&=\frac{\pi}{2}-\int_0^{\pi/2}\frac{\mathrm{d}\theta}{\sin^2\theta+x^2\cos^2\theta}\\&=\frac{\pi}{2}-\int_0^{\pi/2}\frac{\sec^2\theta\mathrm{d}\theta}{\tan^2\theta+x^2}\\&=\frac{\pi}{2}-\frac{1}{x}\arctan\frac{\tan\theta}{x}\Big|_0^{\pi/2}\\&=\frac{\pi}{2}-\frac{\pi}{2x},\end{aligned}$$

which in turn implies that for $x>0$, $x\neq 1$,

$$F'(x)=\frac{2x}{x^2-1}\left(\frac{\pi}{2}-\frac{\pi}{2x}\right)=\frac{\pi}{x+1}.$$

For $x=1$ one sees immediately that $F'(1)=2\int_0^{\pi/2}\cos^2\theta\mathrm{d}\theta=\frac{\pi}{2}$, agreeing with the formula. Now,

$$F'(x)=\frac{\pi}{x+1}\implies F(x)=\pi\log(x+1)+C.$$

Since $F(1)=\int_0^{\pi/2}\log 1\mathrm{d}\theta=0$, we gather that $C=-\pi\log 2$. Finally thus

$$F(x)=\pi\log(x+1)-\pi\log 2=\pi\log\frac{x+1}{2}.$$

◄

Under certain conditions, the interval of integration in the above theorem need not be compact.

186 Example Given that $\int_0^{+\infty}\frac{\sin x}{x}\mathrm{d}x=\frac{\pi}{2}$, compute $\int_0^{+\infty}\frac{\sin^2 x}{x^2}\mathrm{d}x$.

Solution: ▶ *Put* $I(a) = \int_0^{+\infty} \frac{\sin^2 ax}{x^2} \mathrm{d}x$, *with* $a \geq 0$. *Differentiating both sides with respect to* a, *and making the substitution* $u = 2ax$,

$$\begin{aligned} I'(a) &= \int_0^{+\infty} \frac{2x \sin ax \cos ax}{x^2} \mathrm{d}x \\ &= \int_0^{+\infty} \frac{\sin 2ax}{x} \mathrm{d}x \\ &= \int_0^{+\infty} \frac{\sin u}{u} \mathrm{d}u \\ &= \frac{\pi}{2}. \end{aligned}$$

Integrating each side gives

$$I(a) = \frac{\pi}{2} a + C.$$

Since $I(0) = 0$, *we gather that* $C = 0$. *The desired integral is* $I(1) = \frac{\pi}{2}$. ◀

Homework

Problem 2.5.1 Let $f : [0; +\infty[\times]0; +\infty[\to \mathbb{R}$, $f(r, t) = t^n e^{-r^2/4t}$, where n is a constant. Determine n such that

$$\frac{\partial f}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right).$$

Problem 2.5.2 Let

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad f(x, y) = \min(x, y^2).$$

Find $\frac{\partial f(x, y)}{\partial x}$ and $\frac{\partial f(x, y)}{\partial y}$.

Problem 2.5.3 Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ and $g : \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$f(x, y) = \begin{bmatrix} xy^2 \\ x^2 y \end{bmatrix}, \qquad g(x, y, z) = \begin{bmatrix} x - y + 2z \\ xy \end{bmatrix}.$$

Compute $(f \circ g)'(1, 0, 1)$, if at all defined. If undefined, explain. Compute $(g \circ f)'(1, 0)$, if at all defined. If undefined, explain.

Problem 2.5.4 Let $f(x, y) = \begin{bmatrix} xy \\ x + y \end{bmatrix}$ and $g(x, y) = \begin{bmatrix} x - y \\ x^2 y^2 \\ x + y \end{bmatrix}$ Find $(g \circ f)'(0, 1)$.

Problem 2.5.5 Suppose $g : \mathbb{R} \to \mathbb{R}$ is continuous and $a \in \mathbb{R}$ is a constant. Find the partial derivatives with respect to x and y of

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x, y) = \int_a^{x^2 y} g(t)\, \mathrm{d}t.$$

Problem 2.5.6 Given that $\int_0^b \frac{\mathrm{d}x}{x^2 + a^2} = \frac{1}{a} \arctan \frac{b}{a}$, evaluate $\int_0^b \frac{\mathrm{d}x}{(x^2 + a^2)^2}$.

Problem 2.5.7 Prove that

$$\int_0^{+\infty} \frac{\arctan ax - \arctan x}{x} \mathrm{d}x = \frac{\pi}{2} \log \pi.$$

Problem 2.5.8 Assuming that the equation $xy^2 + 3z = \cos z^2$ defines z implicitly as a function of x and y, find $\frac{\partial z}{\partial x}$.

Problem 2.5.9 If $w = e^{uv}$ and $u = r + s$, $v = rs$, determine $\frac{\partial w}{\partial r}$.

Problem 2.5.10 Let z be an implicitly-defined function of x and y through the equation $(x + z)^2 + (y + z)^2 = 8$. Find $\frac{\partial z}{\partial x}$ at $(1, 1, 1)$.

2.6 Gradients and Directional Derivatives

A function

$$f : \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^m \\ \mathbf{x} & \mapsto & f(\mathbf{x}) \end{array}$$

is called a *vector field*. If $m = 1$, it is called a *scalar field*.

187 Definition Let

$$f: \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ \mathbf{x} & \mapsto & f(\mathbf{x}) \end{array}$$

be a scalar field. The *gradient* of f is the vector defined and denoted by

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(\mathbf{x}) \\ \dfrac{\partial f}{\partial x_2}(\mathbf{x}) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(\mathbf{x}) \end{bmatrix}.$$

The *gradient operator* is the operator

$$\nabla = \begin{bmatrix} \dfrac{\partial}{\partial x_1} \\ \dfrac{\partial}{\partial x_2} \\ \vdots \\ \dfrac{\partial}{\partial x_n} \end{bmatrix}.$$

188 Theorem Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}$ be a scalar field, and assume that f is differentiable in A. Let $K \in \mathbb{R}$ be a constant. Then $\nabla f(\mathbf{x})$ is orthogonal to the surface implicitly defined by $f(\mathbf{x}) = K$.

Proof: *Let*

$$\mathbf{c}: \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}^n \\ t & \mapsto & \mathbf{c}(t) \end{array}$$

be a curve lying on this surface. Choose t_0 so that $\mathbf{c}(t_0) = \mathbf{x}$. Then

$$(f \circ \mathbf{c})(t_0) = f(\mathbf{c}(t)) = K,$$

and using the chain rule

$$f'(\mathbf{c}(t_0))\mathbf{c}'(t_0) = 0,$$

which translates to

$$(\nabla f(\mathbf{x})) \bullet (\mathbf{c}'(t_0)) = 0.$$

Since $\mathbf{c}'(t_0)$ is tangent to the surface and its dot product with $\nabla f(\mathbf{x})$ is 0, we conclude that $\nabla f(\mathbf{x})$ is normal to the surface. ❑

☞ *Let θ be the angle between $\nabla f(\mathbf{x})$ and $\mathbf{c}'(t_0)$. Since*

$$|(\nabla f(\mathbf{x})) \bullet (\mathbf{c}'(t_0))| = ||\nabla f(\mathbf{x})|| ||\mathbf{c}'(t_0)|| \cos\theta,$$

$\nabla f(\mathbf{x})$ is the direction in which f is changing the fastest.

189 Example Find a unit vector normal to the surface $x^3 + y^3 + z = 4$ at the point $(1, 1, 2)$.

Solution: ► *Here $f(x, y, z) = x^3 + y^3 + z - 4$ has gradient*

$$\nabla f(x, y, z) = \begin{bmatrix} 3x^2 \\ 3y^2 \\ 1 \end{bmatrix}$$

which at $(1, 1, 2)$ *is* $\begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$. *Normalising this vector we obtain*

$$\begin{bmatrix} \frac{3}{\sqrt{19}} \\ \frac{3}{\sqrt{19}} \\ \frac{1}{\sqrt{19}} \end{bmatrix}.$$

◀

190 Example Find the direction of the greatest rate of increase of $f(x, y, z) = xye^z$ at the point $(2, 1, 2)$.

Solution: ▶ *The direction is that of the gradient vector. Here*

$$\nabla f(x, y, z) = \begin{bmatrix} ye^z \\ xe^z \\ xye^z \end{bmatrix}$$

which at $(2, 1, 2)$ *becomes* $\begin{bmatrix} e^2 \\ 2e^2 \\ 2e^2 \end{bmatrix}$. *Normalising this vector we obtain*

$$\frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

◀

191 Example Let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by

$$f(x, y, z) = x + y^2 - z^2.$$

Find the equation of the tangent plane to f at $(1, 2, 3)$.

Solution: ▶ *A vector normal to the plane is* $\nabla f(1, 2, 3)$. *Now*

$$\nabla f(x, y, z) = \begin{bmatrix} 1 \\ 2y \\ -2z \end{bmatrix}$$

which is

$$\begin{bmatrix} 1 \\ 4 \\ -6 \end{bmatrix}$$

at $(1, 2, 3)$. *The equation of the tangent plane is thus*

$$1(x - 1) + 4(y - 2) - 6(z - 3) = 0,$$

or

$$x + 4y - 6z = -9.$$

◀

192 Definition Let

$$f: \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ \mathbf{x} & \mapsto & f(\mathbf{x}) \end{array}$$

be a vector field with

$$f(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix}.$$

The *divergence* of f is defined and denoted by

$$\operatorname{div} f(\mathbf{x}) = \nabla \bullet f(\mathbf{x}) = \frac{\partial f_1}{\partial x_1}(\mathbf{x}) + \frac{\partial f_2}{\partial x_2}(\mathbf{x}) + \cdots + \frac{\partial f_n}{\partial x_n}(\mathbf{x}).$$

193 Example If $f(x, y, z) = (x^2, y^2, ye^{z^2})$ then

$$\operatorname{div} f(\mathbf{x}) = 2x + 2y + 2yze^{z^2}.$$

194 Definition Let $g_k : \mathbb{R}^n \to \mathbb{R}^n$, $1 \leq k \leq n-2$ be vector fields with $g_i = (g_{i1}, g_{i2}, \ldots, g_{in})$. Then the *curl* of $(g_1, g_2, \ldots, g_{n-2})$

$$\operatorname{curl}(g_1, g_2, \ldots, g_{n-2})(\mathbf{x}) = \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \cdots & \frac{\partial}{\partial x_n} \\ g_{11}(\mathbf{x}) & g_{12}(\mathbf{x}) & \cdots & g_{1n}(\mathbf{x}) \\ g_{21}(\mathbf{x}) & g_{22}(\mathbf{x}) & \cdots & g_{2n}(\mathbf{x}) \\ \vdots & \vdots & \vdots & \vdots \\ g_{(n-2)1}(\mathbf{x}) & g_{(n-2)2}(\mathbf{x}) & \cdots & g_{(n-2)n}(\mathbf{x}) \end{bmatrix}.$$

195 Example If $f(x, y, z) = (x^2, y^2, ye^{z^2})$ then

$$\operatorname{curl} f((\mathbf{x}, \mathbf{y}, \mathbf{z})) = \nabla \times f(x, y, z) = (e^{z^2})\mathbf{i}.$$

196 Example If $f(x, y, z, w) = (e^{xyz}, 0, 0, w^2)$, $g(x, y, z, w) = (0, 0, z, 0)$ then

$$\operatorname{curl}(f, g)(x, y, z, w) = \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} & \frac{\partial}{\partial x_4} \\ e^{xyz} & 0 & 0 & w^2 \\ 0 & 0 & z & 0 \end{bmatrix} = (xz^2e^{xyz})\mathbf{e}_4.$$

197 Definition Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}$ be a scalar field, and assume that f is differentiable in A. Let $\vec{\mathbf{v}} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ be such that $\mathbf{x} + t\vec{\mathbf{v}} \in A$ for sufficiently small $t \in \mathbb{R}$. Then the *directional derivative of f in the direction of $\vec{\mathbf{v}}$ at the point* $\mathbf{x}$ is defined and denoted by

$$\mathscr{D}_{\vec{\mathbf{v}}} f(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\vec{\mathbf{v}}) - f(\mathbf{x})}{t}.$$

☞ *Some authors require that the vector $\vec{\mathbf{v}}$ in definition 197 be a unit vector.*

198 Theorem Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}$ be a scalar field, and assume that f is differentiable in A. Let $\overrightarrow{v} \in \mathbb{R}^n \setminus \{\overrightarrow{0}\}$ be such that $\overrightarrow{x} + t\overrightarrow{v} \in A$ for sufficiently small $t \in \mathbb{R}$. Then the *directional derivative of f in the direction of $\overrightarrow{v}$ at the point $\overrightarrow{x}$* is given by

$$\nabla f(\mathrm{x}) \bullet \overrightarrow{v}.$$

199 Example Find the directional derivative of $f(x, y, z) = x^3 + y^3 - z^2$ in the direction of $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Solution: ▶ *We have*

$$\nabla f(x, y, z) = \begin{bmatrix} 3x^2 \\ 3y^2 \\ -2z \end{bmatrix}$$

and so

$$\nabla f(x, y, z) \bullet \overrightarrow{v} = 3x^2 + 6y^2 - 6z.$$

◀

Homework

Problem 2.6.1 Let $f(x, y, z) = xe^{yz}$. Find

$$(\nabla f)(2, 1, 1).$$

Problem 2.6.2 Let $f(x, y, z) = \begin{bmatrix} xz \\ e^{xy} \\ z \end{bmatrix}$. Find

$$(\nabla \times f)(2, 1, 1).$$

Problem 2.6.3 Find the tangent plane to the surface $\dfrac{x^2}{2} - y^2 - z^2 = 0$ at the point $(2, -1, 1)$.

Problem 2.6.4 Find the point on the surface

$$x^2 + y^2 - 5xy + xz - yz = -3$$

for which the tangent plane is $x - 7y = -6$.

Problem 2.6.5 Find a vector pointing in the direction in which $f(x, y, z) = 3xy - 9xz^2 + y$ increases most rapidly at the point $(1, 1, 0)$.

Problem 2.6.6 Let $\mathscr{D}_{\overrightarrow{u}} f(x, y)$ denote the directional derivative of f at (x, y) in the direction of the unit vector $\overrightarrow{u}$. If $\nabla f(1, 2) = 2\overrightarrow{i} - \overrightarrow{j}$, find $\mathscr{D}_{(\frac{3}{5}, \frac{4}{5})} f(1, 2)$.

Problem 2.6.7 Use a linear approximation of the function $f(x, y) = e^{x \cos 2y}$ at (0, 0) to estimate $f(0.1, 0.2)$.

Problem 2.6.8 Prove that

$$\nabla \bullet (\mathrm{u} \times \mathrm{v}) = \mathrm{v} \bullet (\nabla \times \mathrm{u}) - \mathrm{u} \bullet (\nabla \times \mathrm{v}).$$

Problem 2.6.9 Find the point on the surface

$$2x^2 + xy + y^2 + 4x + 8y - z + 14 = 0$$

for which the tangent plane is $4x + y - z = 0$.

Problem 2.6.10 Let $\phi : \mathbb{R}^3 \to \mathbb{R}$ be a scalar field, and let $\mathrm{U}, \mathrm{V} : \mathbb{R}^3 \to \mathbb{R}^3$ be vector fields. Prove that

1. $\nabla \bullet \phi \mathrm{V} = \phi \nabla \bullet \mathrm{V} + \mathrm{V} \bullet \nabla \phi$
2. $\nabla \times \phi \mathrm{V} = \phi \nabla \times \mathrm{V} + (\nabla \phi) \times \mathrm{V}$
3. $\nabla \times (\nabla \phi) = \overrightarrow{0}$
4. $\nabla \bullet (\nabla \times \mathrm{V}) = 0$
5. $\nabla(\mathrm{U} \bullet \mathrm{V}) = (\mathrm{U} \bullet \nabla)\mathrm{V} + (\mathrm{V} \bullet \nabla)\mathrm{U} + \mathrm{U} \times (\nabla \times \mathrm{V}) + {} + \mathrm{V} \times (\nabla \times \mathrm{U})$

Problem 2.6.11 Find the angles made by the gradient of $f(x, y) = x^{\sqrt{3}} + y$ at the point $(1, 1)$ with the coordinate axes.

2.7 Levi-Civitta and Einstein

☞ *In this section, unless otherwise noted, we are dealing in the space $\mathbb{R}^3$ and so, subscripts are in the set $\{1, 2, 3\}$.*

200 Definition (Einstein's Summation Convention) In any expression containing subscripted variables appearing twice (and only twice) in any term, the subscripted variables are assumed to be summed over.

☞ *In order to emphasise that we are using Einstein's convention, we will enclose any terms under consideration with* $\olessthan \cdot \ogreaterthan$.

201 Example Using Einstein's Summation convention, the dot product of two vectors $\overrightarrow{x} \in \mathbb{R}^n$ and $\overrightarrow{y} \in \mathbb{R}^n$ can be written as

$$\overrightarrow{x} \bullet \overrightarrow{y} = \sum_{i=1}^{n} x_i y_i = \olessthan x_t y_t \ogreaterthan .$$

202 Example Given that a_i, b_j, c_k, d_l are the components of vectors in $\mathbb{R}^3$, $\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}, \overrightarrow{d}$ respectively, what is the meaning of

$$\olessthan a_i b_i c_k d_k \ogreaterthan ?$$

Solution: ► *We have*

$$\olessthan a_i b_i c_k d_k \ogreaterthan = \sum_{i=1}^{3} a_i b_i \olessthan c_k d_k \ogreaterthan = \overrightarrow{a} \bullet \overrightarrow{b} \olessthan c_k d_k \ogreaterthan = \overrightarrow{a} \bullet \overrightarrow{b} \sum_{k=1}^{3} c_k d_k = (\overrightarrow{a} \bullet \overrightarrow{b})(\overrightarrow{c} \bullet \overrightarrow{d}).$$

◄

203 Example Using Einstein's Summation convention, the ij-th entry $(AB)_{ij}$ of the product of two matrices $A \in \mathrm{M}_{m\times n}(\mathbb{R})$ and $B \in \mathrm{M}_{n\times r}(\mathbb{R})$ can be written as

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} = \olessthan A_{it} B_{tj} \ogreaterthan .$$

204 Example Using Einstein's Summation convention, the trace $\mathrm{tr}\,(A)$ of a square matrix $A \in \mathrm{M}_{n\times n}(\mathbb{R})$ is $\mathrm{tr}\,(A) = \sum_{t=1}^{n} A_{tt} = \olessthan A_{tt} \ogreaterthan$.

205 Example Demonstrate, via Einstein's Summation convention, that if A, B are two $n \times n$ matrices, then

$$\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA) .$$

Solution: ► *We have*

$$\mathrm{tr}\,(AB) = \mathrm{tr}\,((AB)_{ij}) = \mathrm{tr}\,(\olessthan A_{ik} B_{kj} \ogreaterthan) = \olessthan \olessthan A_{tk} B_{kt} \ogreaterthan \ogreaterthan,$$

and

$$\mathrm{tr}\,(BA) = \mathrm{tr}\,((BA)_{ij}) = \mathrm{tr}\,(\olessthan B_{ik} A_{kj} \ogreaterthan) = \olessthan \olessthan B_{tk} A_{kt} \ogreaterthan \ogreaterthan,$$

from where the assertion follows, since the indices are dummy variables and can be exchanged. ◄

206 Definition (Kroenecker's Delta) The symbol δ_{ij} is defined as follows:

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

207 Example It is easy to see that $\olessthan \delta_{ik} \delta_{kj} \ogreaterthan = \sum_{k=1}^{3} \delta_{ik} \delta_{kj} = \delta_{ij}$.

208 Example We see that

$$\olessthan \delta_{ij} a_i b_j \ogreaterthan = \sum_{i=1}^{3} \sum_{j=1}^{3} \delta_{ij} a_i b_j = \sum_{k=1} a_k b_k = \overrightarrow{a} \bullet \overrightarrow{b}.$$

Recall that a *permutation* of distinct objects is a reordering of them. The $3! = 6$ permutations of the index set $\{1, 2, 3\}$ can be classified into *even* or *odd*. We start with the identity permutation 123 and say it is even. Now, for any other permutation, we will say that it is even if it takes an even number of transpositions (switching only two elements in one move) to regain the identity permutation, and odd if it takes an odd number of transpositions to regain the identity permutation. Since

$$231 \to 132 \to 123, \quad 312 \to 132 \to 123,$$

the permutations 123 (identity), 231, and 312 are even. Since

$$132 \to 123, \quad 321 \to 123, \quad 213 \to 123,$$

the permutations 132, 321, and 213 are odd.

209 Definition (Levi-Civitta's Alternating Tensor) The symbol ε_{jkl} is defined as follows:

$$\varepsilon_{jkl} = \begin{cases} 0 & \text{if } \{j, k, l\} \neq \{1, 2, 3\} \\ -1 & \text{if } \begin{pmatrix} 1 & 2 & 3 \\ j & k & l \end{pmatrix} \text{ is an odd permutation} \\ +1 & \text{if } \begin{pmatrix} 1 & 2 & 3 \\ j & k & l \end{pmatrix} \text{ is an even permutation} \end{cases}$$

☞ *In particular, if one subindex is repeated we have* $\varepsilon_{rrs} = \varepsilon_{rsr} = \varepsilon_{srr} = 0$. *Also,*

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1, \qquad \varepsilon_{132} = \varepsilon_{321} = \varepsilon_{213} = -1.$$

210 Example Using the Levi-Civitta alternating tensor and Einstein's summation convention, the cross product can also be expressed, if $\overrightarrow{i} = \overrightarrow{e_1}$, $\overrightarrow{j} = \overrightarrow{e_2}$, $\overrightarrow{k} = \overrightarrow{e_3}$, then

$$\overrightarrow{x} \times \overrightarrow{y} = \olessthan \varepsilon_{jkl} (a_k b_l) \overrightarrow{e_j} \ogreaterthan .$$

211 Example If $A = [a_{ij}]$ is a 3×3 matrix, then, using the Levi-Civitta alternating tensor,

$$\det A = \olessthan \varepsilon_{ijk} a_{1i} a_{2j} a_{3k} \ogreaterthan .$$

212 Example Let $\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}$ be vectors in $\mathbb{R}^3$. Then

$$\overrightarrow{x} \bullet (\overrightarrow{y} \times \overrightarrow{z}) = \olessthan x_i (\overrightarrow{y} \times \overrightarrow{z})_i \ogreaterthan = \olessthan x_i \varepsilon_{ikl} (y_k z_l) \ogreaterthan .$$

Homework

Problem 2.7.1 Let $\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}$ be vectors in $\mathbb{R}^3$. Demonstrate that

$$\olessthan x_i y_i z_j \ogreaterthan = (\overrightarrow{x} \bullet \overrightarrow{y}) \overrightarrow{z} .$$

2.8 Extrema

We now turn to the problem of finding maxima and minima for vector functions. As in the one-variable case, the derivative will provide us with information about the extrema, and the "second derivative" will provide us with information about the nature of these extreme points.

To define an analogue for the second derivative, let us consider the following. Let $A \subset \mathbb{R}^n$ and $f : A \to \mathbb{R}^m$ be differentiable on A. We know that for fixed $\mathbf{x}_0 \in A$, $\mathscr{D}_{\mathbf{x}_0}(f)$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. This means that we have a function

$$T : \begin{array}{ccc} A & \to & \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m) \\ \mathbf{x} & \mapsto & \mathscr{D}_{\mathbf{x}}(f) \end{array},$$

where $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ denotes the space of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. Hence, if we differentiate T at $\mathbf{x}_0$ again, we obtain a linear transformation $\mathscr{D}_{\mathbf{x}_0}(T) = \mathscr{D}_{\mathbf{x}_0}(\mathscr{D}_{\mathbf{x}_0}(f)) = \mathscr{D}^2_{\mathbf{x}_0}(f)$ from $\mathbb{R}^n$ to $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$. Hence, given $\mathbf{x}_1 \in \mathbb{R}^n$, we have $\mathscr{D}^2_{\mathbf{x}_0}(f)(\mathbf{x}_1) \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$. Again, this means that given $\mathbf{x}_2 \in \mathbb{R}^n$, $\mathscr{D}^2_{\mathbf{x}_0}(f)(\mathbf{x}_1))(\mathbf{x}_2) \in \mathbb{R}^m$. Thus the function

$$B_{\mathbf{x}_0} : \begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \to & \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m) \\ (\mathbf{x}_1, \mathbf{x}_2) & \mapsto & \mathscr{D}^2_{\mathbf{x}_0}(f)(\mathbf{x}_1, \mathbf{x}_2) \end{array}$$

is well defined, and linear in each variable $\mathbf{x}_1$ and $\mathbf{x}_2$, that is, it is a *bilinear* function.

Just as the Jacobi matrix was a handy tool for finding a matrix representation of $\mathscr{D}_{\mathbf{x}}(f)$ in the natural bases, when f maps into $\mathbb{R}$, we have the following analogue representation of the second derivative.

213 Theorem Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \to \mathbb{R}$ be twice differentiable on A. Then the matrix of $\mathscr{D}^2_{\mathbf{x}}(f) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with respect to the standard basis is given by the *Hessian matrix:*

$$\mathscr{H}_x f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}) \end{bmatrix}$$

214 Example Let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by

$$f(x, y, z) = xy^2z^3.$$

Then

$$\mathscr{H}_{(x,y,z)} f = \begin{bmatrix} 0 & 2yz^3 & 3y^2z^2 \\ 2yz^3 & 2xz^3 & 6xyz^2 \\ 3y^2z^2 & 6xyz^2 & 6xy^2z \end{bmatrix}$$

From the preceding example, we notice that the Hessian is symmetric, as the mixed partial derivatives $\frac{\partial^2}{\partial x \partial y} f = \frac{\partial^2}{\partial y \partial x} f$, etc., are equal. This is no coincidence, as guaranteed by the following theorem.

215 Theorem Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \to \mathbb{R}$ be twice differentiable on A. If $\mathscr{D}^2_{\mathbf{x}_0}(f)$ is continuous, then $\mathscr{D}^2_{\mathbf{x}_0}(f)$ is symmetric, that is, $\forall (\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{R}^n \times \mathbb{R}^n$ we have

$$\mathscr{D}^2_{\mathbf{x}_0}(f)(\mathbf{x}_1, \mathbf{x}_2) = \mathscr{D}^2_{\mathbf{x}_0}(f)(\mathbf{x}_2, \mathbf{x}_1).$$

We are now ready to study extrema in several variables. The basic theorems resemble those of one-variable calculus. First, we make some analogous definitions.

216 Definition Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \to \mathbb{R}$. If there is some open ball $B_{\mathbf{x}_0}(\mathbf{r})$ on which $\forall \mathbf{x} \in B_{x_0}(\mathbf{r}), \quad f(\mathbf{x}_0) \geq f(\mathbf{x})$, we say that $f(\mathbf{x}_0)$ is a *local maximum* of f. Similarly, if there is some open ball $B_{\mathbf{x}_1}(\mathbf{r})$ on which $\forall \mathbf{x} \in B_{\mathbf{x}_0}(\mathbf{r}'), \quad f(\mathbf{x}_1) \leq f(\mathbf{x})$, we say that $f(\mathbf{x}_1)$ is a *local maximum* of f. A point is called an *extreme point* if it is either a local minimum or local maximum. A point $\mathbf{t}$ is called a *critical point* if f is differentiable at $\mathbf{t}$ and $\mathscr{D}_{\mathbf{t}}(f) = \mathbf{0}$. A critical point which is neither a maxima nor a minima is called a *saddle point*.

217 Theorem Let $A \subseteq \mathbb{R}^n$ be an open set, and $f : A \to \mathbb{R}$ be differentiable on A. If $\mathbf{x}_0$ is an extreme point, then $\mathscr{D}_{\mathbf{x}_0}(f) = \mathbf{0}$, that is, $\mathbf{x}_0$ is a critical point. Moreover, if f is twice-differentiable with continuous second derivative and $\mathbf{x}_0$ is a critical point such that $\mathscr{H}_{\mathbf{x}_0} f$ is negative definite, then f has a local maximum at $\mathbf{x}_0$. If $\mathscr{H}_{\mathbf{x}_0} f$ is positive definite, then f has a local minimum at $\mathbf{x}_0$. If $\mathscr{H}_{\mathbf{x}_0} f$ is indefinite, then f has a saddle point. If $\mathscr{H}_{\mathbf{x}_0} f$ is semi-definite (positive or negative), the test is inconclusive.

218 Example Find the critical points of

$$f : \begin{array}{lll} \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & x^2 + xy + y^2 + 2x + 3y \end{array}$$

and investigate their nature.

Solution: ► *We have*

$$(\nabla f)(x, y) = \begin{bmatrix} 2x + y + 2 \\ x + 2y + 3 \end{bmatrix},$$

and so to find the critical points we solve

$$2x + y + 2 = 0,$$

$$x + 2y + 3 = 0,$$

which yields $x = -\dfrac{1}{3}, y = -\dfrac{4}{3}$. *Now,*

$$\mathscr{H}_{(x,y)} f = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix},$$

which is positive definite, since $\Delta_1 = 2 > 0$ *and* $\Delta_2 = \det \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = 3 > 0$. *Thus* $\mathbf{x}_0 = \left(-\dfrac{1}{3}, -\dfrac{4}{3}\right)$ *is a relative minimum and we have*

$$-\frac{7}{3} = f\left(-\frac{1}{3}, -\frac{4}{3}\right) \leq f(x, y) = x^2 + xy + y^2 + 2x + 3y.$$

◄

219 Example Find the extrema of

$$f : \begin{array}{lll} \mathbb{R}^3 & \to & \mathbb{R} \\ (x, y, z) & \mapsto & x^2 + y^2 + 3z^2 - xy + 2xz + yz \end{array}.$$

Solution: ► *We have*

$$(\nabla f)(x,y,z) = \begin{bmatrix} 2x - y + 2z \\ 2y - x + z \\ 6z + 2x + y \end{bmatrix},$$

which vanishes when $x = y = z = 0$. *Now,*

$$\mathscr{H}_r f = \begin{bmatrix} 2 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & 1 & 6 \end{bmatrix},$$

which is positive definite, since $\Delta_1 = 2 > 0$, $\Delta_2 = \det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 3 > 0$, *and* $\Delta_3 = \det \begin{bmatrix} 2 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & 1 & 6 \end{bmatrix} = 4 > 0$. *Thus f has a relative minimum at* $(0,0,0)$ *and*

$$0 = f(0,0,0) \leq f(x,y,z) = x^2 + y^2 + 3z^2 - xy + 2xz + yz.$$

◄

220 Example Let $f(x,y) = x^3 - y^3 + axy$, with $a \in \mathbb{R}$ a parameter. Determine the nature of the critical points of f.

Solution: ► *We have*

$$(\nabla f)(x,y) = \begin{bmatrix} 3x^2 + ay \\ -3y^2 + ax \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies 3x^2 = -ay, \quad 3y^2 = ax.$$

If $a = 0$, *then* $x = y = 0$ *and so* $(0,0)$ *is a critical point. If* $a \neq 0$ *then*

$$\begin{aligned} 3\left(3\frac{y^2}{a}\right)^2 = -ay \quad & \implies \quad 27y^4 = -a^3 y \\ & \implies \quad y(27y^3 + a^3) = 0 \\ & \implies \quad y(3y + a)(9y^2 - 3ay + a^2) = 0 \\ & \implies \quad y = 0 \quad \text{or} \quad y = -\frac{a}{3}. \end{aligned}$$

If $y = 0$ *then* $x = 0$, *so again* $(0,0)$ *is a critical point. If* $y = -\frac{a}{3}$ *then* $x = \frac{3}{a} \cdot \left(-\frac{a}{3}\right)^2 = \frac{a}{3}$ *so* $\left(\frac{a}{3}, -\frac{a}{3}\right)$ *is a critical point.*
Now,

$$\mathcal{H}_{f(x,y)} = \begin{bmatrix} 6x & a \\ a & -6y \end{bmatrix} \implies \Delta_1 = 6x, \quad \Delta_2 = -36xy - a^2.$$

At $(0,0)$, $\Delta_1 = 0$, $\Delta_2 = -a^2$. *If* $a \neq 0$ *then there is a saddle point. At* $\left(\frac{a}{3}, -\frac{a}{3}\right)$, $\Delta_1 = 2a$, $\Delta_2 = 3a^2$, *whence* $\left(\frac{a}{3}, -\frac{a}{3}\right)$ *will be a local minimum if* $a > 0$ *and a local maximum if* $a < 0$. ◄

Homework

Problem 2.8.1 Determine the nature of the critical points of $f(x, y) = y^2 - 2x^2y + 4x^3 + 20x^2$.

Problem 2.8.2 Determine the nature of the critical points of $f(x, y) = (x - 2)^2 + 2y^2$.

Problem 2.8.3 Determine the nature of the critical points of $f(x, y) = (x - 2)^2 - 2y^2$.

Problem 2.8.4 Determine the nature of the critical points of $f(x, y) = x^4 + 4xy - 2y^2$.

Problem 2.8.5 Determine the nature of the critical points of $f(x, y) = x^4 + y^4 - 2x^2 + 4xy - 2y^2$.

Problem 2.8.6 Determine the nature of the critical points of $f(x, y, z) = x^2 + y^2 + z^2 - xy + x - 2z$.

Problem 2.8.7 Determine the nature of the critical points of $f(x, y) = x^4 + y^4 - 2(x - y)^2$.

Problem 2.8.8 Determine the nature of the critical points of

$$f(x, y, z) = 4x^2z - 2xy - 4x^2 - z^2 + y.$$

Problem 2.8.9 Find the extrema of

$$f(x, y, z) = x^2 + y^2 + z^2 + xyz.$$

Problem 2.8.10 Find the extrema of $f(x, y, z) = x^2y + y^2z + 2x - z$.

Problem 2.8.11 Determine the nature of the critical points of

$$f(x, y, z) = 4xyz - x^4 - y^4 - z^4.$$

Problem 2.8.12 Determine the nature of the critical points of $f(x, y, z) = xyz(4 - x - y - z)$.

Problem 2.8.13 Determine the nature of the critical points of

$$g(x, y, z) = xyze^{-x^2-y^2-z^2}.$$

Problem 2.8.14 Let $f(x, y) = \int_{y^2-x}^{x^2+y} g(t)\mathrm{d}t$, where g is a continuously differentiable function defined over all real numbers and $g(0) = 0, g'(0) \neq 0$. Prove that $(0, 0)$ is a saddle point for f.

Problem 2.8.15 Find the minimum of

$$F(x, y) = (x - y)^2 + \left(\frac{\sqrt{144 - 16x^2}}{3} - \sqrt{4 - y^2}\right)^2,$$

for $-3 \leq x \leq 3$, $-2 \leq y \leq 2$.

2.9 Lagrange Multipliers

In some situations we wish to optimise a function given a set of constraints. For such cases, we have the following.

221 Theorem Let $A \subseteq \mathbb{R}^n$ and let $f : A \to \mathbb{R}$, $g : A \to \mathbb{R}$ be functions whose respective derivatives are continuous. Let $g(\mathrm{x}_0) = c_0$ and let $S = g^{-1}(c_0)$ be the level set for g with value c_0, and assume $\nabla g(\mathrm{x}_0) \neq 0$. If the restriction of f to S has an extreme point at x_0, then $\exists \lambda \in \mathbb{R}$ such that

$$\nabla f(\mathrm{x}_0) = \lambda \nabla g(\mathrm{x}_0).$$

☞ *The above theorem only locates extrema, it does not say anything concerning the nature of the critical points found.*

222 Example Optimise $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 - y^2$ given that $x^2 + y^2 = 1$.

Solution: ► *Let* $g(x, y) = x^2 + y^2 - 1$. *We solve*

$$\nabla f \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \nabla g \begin{bmatrix} x \\ y \end{bmatrix}$$

for x, y, λ. *This requires*

$$\begin{bmatrix} 2x \\ -2y \end{bmatrix} = \begin{bmatrix} 2x\lambda \\ 2y\lambda \end{bmatrix}.$$

From $2x = 2x\lambda$ *we get either* $x = 0$ *or* $\lambda = 1$. *If* $x = 0$ *then* $y = \pm 1$ *and* $\lambda = -1$. *If* $\lambda = 1$, *then* $y = 0, x = \pm 1$. *Thus the potential critical points are* $(\pm 1, 0)$ *and* $(0, \pm 1)$. *If* $x^2 + y^2 = 1$ *then*

$$f(x,y) = x^2 - (1 - x^2) = 2x^2 - 1 \geq -1,$$

and

$$f(x,y) = 1 - y^2 - y^2 = 1 - 2y^2 \leq 1.$$

Thus $(\pm 1, 0)$ *are maximum points and* $(0, \pm 1)$ *are minimum points.* ◀

223 Example Find the maximum and the minimum points of $f(x,y) = 4x+3y$, subject to the constraint $x^2 + 4y^2 = 4$, using Lagrange multipliers.

Solution: ► *Putting* $g(x,y) = x^2 + 4y^2 - 4$ *we have*

$$\nabla f(x,y) = \lambda \nabla g(x,y) \implies \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 8y \end{bmatrix}.$$

Thus $4 = 2\lambda x, \quad 3 = 8\lambda y$. *Clearly then* $\lambda \neq 0$. *Upon division we find* $\frac{x}{y} = \frac{16}{3}$. *Hence*

$$x^2 + 4y^2 = 4 \implies \frac{256}{9}y^2 + 4y^2 = 4 \implies y = \pm\frac{3}{\sqrt{73}}, x = \pm\frac{16}{\sqrt{73}}.$$

The maximum is clearly then

$$4\left(\frac{16}{\sqrt{73}}\right) + 3\left(\frac{3}{\sqrt{73}}\right) = \sqrt{73},$$

and the minimum is $-\sqrt{73}$. ◀

224 Example Let $a > 0, b > 0, c > 0$. Determine the maximum and minimum values of $f(x,y,z) = \frac{x}{a} + \frac{y}{b} + \frac{z}{c}$ on the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

Solution: ► *We use Lagrange multipliers. Put* $g(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1$. *Then*

$$\nabla f(x,y,z) = \lambda \nabla g(x,y,z) \iff \begin{bmatrix} 1/a \\ 1/b \\ 1/c \end{bmatrix} = \lambda \begin{bmatrix} 2x/a^2 \\ 2y/b^2 \\ 2z/c^2 \end{bmatrix}.$$

It follows that $\lambda \neq 0$. *Hence* $x = \frac{a}{2\lambda}, y = \frac{b}{2\lambda}, z = \frac{c}{2\lambda}$. *Since* $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$, *we deduce* $\frac{3}{4\lambda^2} = 1$ *or* $\lambda = \pm\frac{\sqrt{3}}{2}$. *Since* a, b, c *are positive,* f *will have a maximum when all* x, y, z *are positive and a minimum when all* x, y, z *are negative. Thus the maximum is when*

$$x = \frac{a}{\sqrt{3}}, y = \frac{b}{\sqrt{3}}, z = \frac{c}{\sqrt{3}},$$

and

$$f(x,y,z) \leq \frac{3}{\sqrt{3}} = \sqrt{3}$$

and the minimum is when

$$x = -\frac{a}{\sqrt{3}}, y = -\frac{b}{\sqrt{3}}, z = -\frac{c}{\sqrt{3}},$$

and

$$f(x,y,z) \geq -\frac{3}{\sqrt{3}} = -\sqrt{3}.$$

Aliter: *Using the CBS Inequality,*

$$\left|\frac{x}{a}\cdot 1+\frac{y}{b}\cdot 1+\frac{z}{c}\cdot 1\right| \leq \left(\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\right)^{1/2}\left(1^2+1^2+1^2\right)^{1/2} = (1)\sqrt{3} \implies -\sqrt{3} \leq \frac{x}{a}+\frac{y}{b}+\frac{z}{c} \leq \sqrt{3}.$$

◀

225 Example Let $a > 0, b > 0, c > 0$. Determine the maximum volume of the parallelepiped with sides parallel to the axes that can be enclosed inside the ellipsoid $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} = 1$.

Solution: ▶ *Let $2x$, $2y$, $2z$, be the dimensions of the box. We must maximise $f(x,y,z) = 8xyz$ subject to the constraint $g(x,y,z) = \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}-1$. Using Lagrange multipliers,*

$$\nabla f(x,y,z) = \lambda\nabla g(x,y,z) \iff \begin{bmatrix} 8yz \\ 8xz \\ 8xy \end{bmatrix} = \lambda\begin{bmatrix} 2x/a^2 \\ 2y/b^2 \\ 2z/c^2 \end{bmatrix} \implies 4yz = \lambda\frac{x}{a^2},\ 4xz = \lambda\frac{y}{b^2},\ 4xy = \lambda\frac{z}{c^2}.$$

Multiplying the first inequality by x, the second by y, the third by z, and adding,

$$4xyz = \lambda\frac{x^2}{a^2},\ 4xyz = \lambda\frac{y^2}{b^2},\ 4xyz = \lambda\frac{z^2}{c^2}, \implies 12xyz = \lambda\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\right) = \lambda.$$

Hence

$$\frac{\lambda}{3} = \lambda\frac{x^2}{a^2} = \lambda\frac{y^2}{b^2} = \lambda\frac{z^2}{c^2}.$$

If $\lambda = 0$, then $8xyz = 0$, which minimises the volume. If $\lambda \neq 0$, then

$$x = \frac{a}{\sqrt{3}}, \qquad y = \frac{b}{\sqrt{3}}, \qquad z = \frac{c}{\sqrt{3}},$$

and the maximum value is

$$8xyz \leq 8\frac{abc}{3\sqrt{3}}.$$

Aliter: *Using the AM-GM Inequality,*

$$(x^2y^2z^2)^{1/3} = (abc)^{2/3}\left(\frac{x^2}{a^2}\cdot\frac{y^2}{b^2}\cdot\frac{z^2}{a^2}\right)^{1/3} \leq (abc)^{2/3}.\frac{\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}}{3} = \frac{1}{3} \implies 8xyz \leq \frac{8}{3\sqrt{3}}(abc).$$

◀

Homework

Problem 2.9.1 A closed box (with six outer faces), has fixed surface area of S square units. Find its maximum volume using Lagrange multipliers. That is, subject to the constraint $2ab+2bc+2ca = S$, you must maximise abc.

Problem 2.9.2 Consider the problem of finding the closest point P' on the plane $\Pi : ax+by+cz = d$,

a, b, c non-zero constants with $a+b+c \neq d$ to the point $P(1,1,1)$. In this problem, you will do this in three essentially different ways.

1. Do this by a geometric argument, arguing the the point P' closest to P on Π is on the perpendicular passing through P and P'.
2. Do this by means of Lagrange multipliers, by minimising a suitable function $f(x,y,z)$ subject to the constraint $g(x,y,z) = ax + by + cz - d$.
3. Do this considering the unconstrained extrema of a suitable function $h\left(x, y, \dfrac{d - ax - by}{c}\right)$.

Problem 2.9.3 Given that x, y are positive real numbers such that $x^4 + 81y^4 = 36$ find the maximum of $x + 3y$.

Problem 2.9.4 If x, y, z are positive real numbers such that $x^2y^3z = \dfrac{1}{6^2}$, what is the minimum value of $f(x,y,z) = 2x + 3y + z$?

Problem 2.9.5 Find the maximum and the minimum values of $f(x,y) = x^2 + y^2$ subject to the constraint $5x^2 + 6xy + 5y^2 = 8$.

Problem 2.9.6 Let $a > 0, b > 0, p > 1$. Maximise $f(x,y) = ax + by$ subject to the constraint $x^p + y^p = 1$.

Problem 2.9.7 Find the extrema of

$$f(x,y,z) = x^2 + y^2 + z^2$$

subject to the constraint

$$(x-1)^2 + (y-2)^2 + (z-3)^2 = 4.$$

Problem 2.9.8 Find the axes of the ellipse

$$5x^2 + 8xy + 5y^2 = 9.$$

Problem 2.9.9 Optimise $f(x,y,z) = x + y + z$ subject to $x^2 + y^2 = 2$, and $x + z = 1$.

Problem 2.9.10 Let x, y be strictly positive real numbers with $x + y = 1$. What is the maximum value of $x + \sqrt{xy}$?

Problem 2.9.11 Let a, b be positive real constants. Maximise $f(x,y) = x^a e^{-x} y^b e^{-y}$ on the triangle in $\mathbb{R}^2$ bounded by the lines $x \geq 0$, $y \geq 0$, $x + y \leq 1$.

Problem 2.9.12 Determine the extrema of $f(x,y) = \cos^2 x + \cos^2 y$ subject to the constraint $x - y = \dfrac{\pi}{4}$.

Problem 2.9.13 Determine the extrema of $f(x,y,z) = x - 2y + 2z$ subject to the constraint $x^2 + y^2 + z^2 = 1$.

Problem 2.9.14 Find the points on the curve determined by the equations

$$x^2 + xy + y^2 - z^2 = 1, \quad x^2 + y^2 = 1$$

which are closest to the origin.

Problem 2.9.15 Does there exist a polynomial in two variables with real coefficients $p(x,y)$ such that $p(x,y) > 0$ for all x and y and that for all real numbers $c > 0$ there exists $(x_0, y_0) \in \mathbb{R}^2$ such that $p(x_0, y_0) = c$?

Problem 2.9.16 Maximise

$$f(x,y,z) = \log x + \log y + 3 \log z$$

on the portion of sphere $x^2 + y^2 + z^2 = 5r^2$ which lies on the first octant. Demonstrate using this that for any positive real numbers a, b and c, there follows the inequality

$$abc^3 \leq 27 \left(\frac{a+b+c}{5}\right)^5.$$